A Generalization of the Aumann-Shapley Value for Risk Capital Allocation Problems

Boonen, T.J.; De Waegenaere, A.M.B.; Norde, H.W.

Publication date:
2012

CentER

Discussion Paper

No. 2012-091

A GENERALIZATION OF THE AUMANN-SHAPLEY VALUE FOR RISK ALLOCATION PROBLEMS

By

Tim Boonen, Anja De Waegenaere,
Henk Norde

28 November, 2012

ISSN 0924-7815
ISSN 2213-9532

TILBURG UNIVERSITY

A generalization of the Aumann-Shapley value for risk capital allocation problems

Tim Boonen* Anja De Waegenaere† Henk Norde‡

November 22, 2012

Abstract

This paper analyzes risk capital allocation problems. For risk capital allocation problems, the aim is to allocate the risk capital of a firm to its divisions. Risk capital allocation is of central importance in risk-based performance measurement. We consider a case in which the aggregate risk capital is determined via a coherent risk measure. The academic literature advocates an allocation rule that, in game-theoretic terms, is equivalent to using the Aumann-Shapley value as solution concept. This value is however not well-defined in case a differentiability condition is not satisfied. As an alternative, we introduce an allocation rule inspired by the Shapley value in a fuzzy setting. We take a grid on a fuzzy participation set, define paths on this grid and construct an allocation rule based on a path. Then, we define a rule as the limit of the average over these allocations, when the grid size converges to zero. We introduce this rule for a broad class of coherent risk measures. We show that if the Aumann-Shapley value is well-defined, the allocation rule coincides with it. If the Aumann-Shapley value is not defined, which is due to non-differentiability problems, the allocation rule specifies an explicit allocation. It corresponds with the Mertens value, which is originally characterized in an axiomatic way (Mertens, 1988), whereas we provide an asymptotic argument.

JEL-Classification: C71, G32

1 Introduction

This paper proposes a rule to allocate *risk capital* among divisions within a firm. Regulators require that financial institutions withhold a level of capital that is invested safely in order to mitigate the effects of adverse events such as, for example, a financial crisis. This amount of capital is referred to as risk capital. Regulatory requirements focus at the level of risk capital to be withheld at firm level. Our focus is on how this amount of risk capital is allocated to different business divisions within the firm.[1] This problem is called the *risk capital allocation problem.*

There are several reasons why firms want to allocate risk capital to divisions. First, allocating risk capital is important for performance evaluation. Investment activities of financial institutions are typically divided into different portfolios, with different divisions within the firm being responsible for different portfolios. It is not uncommon that the managers of these divisions are evaluated on the basis of the return earned on the amount of risk capital to be withheld for their portfolio. This requires an allocation of risk capital to divisions that is perceived as "fair" by the managers. Second, allocating risk capital to business divisions is important for decisions regarding whether to increase or decrease the engagement in the activities of certain divisions. The attractiveness of a specific risky activity (e.g., a specific financial investment) is typically evaluated by means of a risk-return trade-off. Evaluating the performance of a division's activities in isolation, however, can be very misleading. For example, the activity might seem highly risky in isolation, but may be useful in hedging risk in other divisions's

*Corresponding author: Tel: +31 13 4662373, Fax: +31 13 4663280, Email address: t.j.boonen@uvt.nl, Department of Econometrics and OR, Tilburg University, CentER for Economic Research and Netspar, PO box 90153, 5000 LE Tilburg.

†Department of Accountancy and Department of Econometrics and OR, Tilburg University, CentER for Economic Research and Netspar.

‡Department of Econometrics and OR, Tilburg University, CentER for Economic Research.

[1] Alternatively, one can interpret a division as a financial portfolio.

activities.[2] One approach to evaluate the attractiveness of increasing the engagement in the activities of a specific division taking into account potential hedge effects is to determine the effect of increasing the level of the activities on the allocation of risk capital to all divisions.

The allocation problem is non-trivial because whenever a coherent risk measure (Artzner, Delbaen, Eber and Heath, 1999) is used to determine risk capital, the amount of risk capital to be withheld for the firm as a whole would typically be lower than the sum of the amounts of risk capital that would need to be withheld for each division in isolation. The reason is that the individual risks associated with the divisions are typically not perfectly correlated, and, hence, there can be some hedge potential from combining the risks. The allocation rule then determines how the benefits of this hedge potential are allocated to the divisions.

There is a large literature on capital allocation rules, with approaches based on finance (e.g., Tasche, 1999), optimization (e.g., Dhaene et al., 2003; Laeven and Goovaerts, 2006) and game theory (e.g., Denault, 2001; Tsanakas and Barnett, 2003; Tsanakas, 2004 and 2009). Our focus in this paper is on game-theoretic approaches to allocating risk capital. A game-theoretic approach that has received considerable attention is the one of Denault (2001). He models the risk capital allocation problem as a *fuzzy game*. Specifically, he defines a risk capital allocation function as a function that assigns an amount of risk capital to every collection of fractions of divisions. The fraction of a division included in a collection is referred to as the *participation level* of that division. He then considers risk capital allocations that satisfy the stability condition that requires that, for any given collection of fractions of divisions, the amount of risk capital allocated to that collection is weakly lower than the amount of risk capital that they would need to withhold if they would separate from the firm. In game-theoretic terms, this condition means that the allocation is an element of the *fuzzy core*. Denault specifies a number of other desirable properties of a risk capital allocation rule, and shows that the *Aumann-Shapley value* (Aumann and Shapley, 1974) is the only allocation rule that is in the fuzzy core and satisfies these additional properties.[3] Moreover, Kalkbrener (2005) imposes a diversification axiom that requires the risk capital allocation of a division not to exceed its corresponding stand-alone risk capital. The Aumann-Shapley value is then characterized as the only allocation rule that satisfies this condition and two more technical conditions.

The Aumann-Shapley value as a risk capital allocation rule has received considerable attention in the literature. Financial and economic arguments in favor of the Aumann-Shapley value are provided by, e.g., Tasche (1999) and Myers and Read (2001). One of the drawbacks, however, of the Aumann-Shapley value is that it requires *partial differentiability* of the fuzzy risk capital allocation function at the level of full participation of each division. It is well-known that the fuzzy risk capital function is generally not differentiable everywhere when the probability distributions of the risks associated with the divisions are not continuous (see, e.g., Tasche, 1999). We propose a generalization of the Aumann-Shapley value that is well-defined even when the risk capital function is not differentiable. The rule that we propose is inspired by the idea underlying the *Shapley value* (Shapley, 1953) for non-fuzzy cooperative games. We first discretize the participation levels of divisions by considering a finite grid of participation levels. Then, for any given discrete path on the grid starting from no participation (the participation profile where the participation level of each division is zero) and ending at full participation (the participation profile with full participation of each division), we determine the corresponding path-based allocation. Specifically, in each step of the path, the participation level of exactly one division is increased, and the corresponding difference in risk capital is allocated to that division. Proceeding in this way along the path, the total risk capital will be allocated once the path reaches the level of full participation. This procedure yields a risk capital allocation for every possible path. Moreover, the average of the corresponding risk capital allocations over all possible paths is also a risk capital allocation.[4] We show that when the grid size converges to zero, this average converges as well. The allocation rule that we propose in this paper equals this asymptotic value. We refer to it as the *Weighted-Aumann Shapley* value. For risk capital allocation problems for which the corresponding risk capital function is differentiable at the level of full participation, the Weighted-Aumann Shapley value coincides with the Aumann-Shapley value. In contrast to the Aumann-Shapley value, however, the Weighted-Aumann Shapley value is well-defined even when the risk capital allocation function

[2]An example would be an insurance company that holds both annuities and death benefit insurance. Both types of liabilities are sensitive to longevity risk (the risk associated with unpredictable changes in survival rates in a population). In isolation, each of these liabilities could be evaluated as relatively risky. However, the death benefit insurance provides hedge potential for the annuity portfolio. Van Gulick, De Waegenaere and Norde (2012) show the impact of this hedge potential on the allocation of risk capital.

[3]For general production functions, the Aumann-Shapley value is characterized by, e.g., Aumann and Shapley (1974), Aubin (1981), Billera and Heath (1982) and Mirman and Tauman (1982) .

[4]The construction of a rule as average over paths is in line with, e.g., Moulin (1995) and Sprumont (2005), who both consider a discrete production problem.

is non-differentiable.

In the seminal book of Aumann and Shapley (1974) the Aumann-Shapley value is introduced and characterized for special classes of games with a continuum of players. Roughly speaking, the characteristic functions of these games are obtained as differentiable function of a finite number of non-atomic probability measures. Aumann and Shapley moreover provide the well-known "diagonal formula" for their value. Mertens (1980 and 1988) extends the Aumann-Shapley value and its axiomatic characterization to a much larger class of vector measure games by dropping the differentiability assumption. An overview of the Mertens value for vector measure games is given by Neyman (2002). Since fuzzy games can be considered as special examples of vector measure games the Aumann Shapley value can be computed for these games as well under some differentiability assumptions and the Mertens value under much milder assumptions. In fact, Mertens shows that in order to compute the Mertens value the Aumann-Shapley diagonal formula should be generalized to a diagonal formula where an expectation of partial derivatives along random, small perturbations around the diagonal should be integrated.

Aumann and Shapley (1974) show that under very strong assumptions their value (and hence the Mertens value) can be obtained via an asymptotic approach. However, in Example 19.2 of their book they show that fuzzy games, corresponding to convex, piecewise affine functions (like the fuzzy games related to risk capital allocation problems which we consider in this paper), do not satisfy this strong assumption (also pointed out by Neyman and Smorodinsky, 2004). In this paper we provide an allocation rule that follows a much weaker asymptotic approach than the one used by Aumann and Shapley (1974). In return, we get that our approach is convergent for all fuzzy games related to risk capital allocation problems. Moreover, our value happens to coincide with the Mertens value. By the way, an axiomatization of the Mertens value on the class of piece-wise linear fuzzy games is provided by Haimanko (2001).

We also show that the corresponding risk capital allocation rule satisfies a number of desirable properties. Some of these properties are known to be satisfied by the regular Aumann-Shapley value on the class of risk capital allocation problems for which the Aumann-Shapley value is well-defined. Moreover, the approach that we use to characterize the allocation rule allows us to give an explicit formula for the corresponding capital allocations. The specific formula has a geometric interpretation.

This paper is set out as follows. In Section 2, we define risk capital and risk capital allocation problems. Two of the most prominent game-theoretic solution concepts for allocation problems are discussed in Section 3, namely the Shapley value and the Aumann-Shapley value. In Section 4, we provide the structure of the risk capital function. In Section 5 we define a class of path-based allocation rules. In Section 6, we introduce an allocation rule based on the average of path-based allocations. We show that the corresponding allocation rule can be seen as a generalization of the Aumann-Shapley value, and that it satisfies some desirable properties. We also provide a closed form expression with a geometric interpretation.

2 Risk measures and risk capital allocation problems

In this paper, we propose a rule to allocate risk capital among divisions. The firm uses a risk measure to determine this capital. In this section, we briefly introduce risk measures and risk capital allocation problems.

2.1 Risk measures

In this subsection, we discuss risk measures as in Artzner et al. (1999) and Delbaen (2000). Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a fixed probability space, i.e., Ω is the state space, $\mathcal{F}$ is a σ-algebra on Ω and $\mathbb{P}$ is the physical probability measure on $(\Omega, \mathcal{F})$. We denote $\mathcal{P}(\Omega, \mathcal{F})$ as the set of all probability measures on $(\Omega, \mathcal{F})$ and $L^\infty(\Omega, \mathcal{F}, \mathbb{P})$ for the space of all bounded, measurable, real valued stochastic variables. If there is no confusion possible, we write $L^\infty = L^\infty(\Omega, \mathcal{F}, \mathbb{P})$. We interpret a realization of a stochastic variable as a future loss.

A risk measure is a function $\rho : L^\infty \to \mathbb{R}$.[5] So, a risk measure maps stochastic variables into real numbers. It serves as a measure to determine the cash reserve for holding risk. The purpose of this reserve is to make the risk acceptable to the regulator. In this paper, we only focus on *coherent* risk measures. Coherence is first introduced by Artzner et al. (1999). A risk measure ρ is called coherent if it satisfies the following four properties:

[5] Here, we assume that ρ is only defined on L^∞. For a discussion of risk measures on the class of all stochastic variables, we refer to Delbaen (2000).

- *Sub-additivity*: For all $X, Y \in L^\infty$, we have

$$\rho(X+Y) \leq \rho(X) + \rho(Y).$$

- *Monotonicity*: For all $X, Y \in L^\infty$ such that $X(\omega) \geq Y(\omega)$ holds almost surely for $\omega \in \Omega$ with respect to the measure $\mathbb{P}$, we have

$$\rho(X) \geq \rho(Y).$$

- *Positive Homogeneity*: For every $X \in L^\infty$ and every $c > 0$, we have

$$\rho(c \cdot X) = c \cdot \rho(X).$$

- *Translation Invariance*: For every $X \in L^\infty$ and every $c \in \mathbb{R}$, we have

$$\rho(X + c \cdot e_\Omega) = \rho(X) + c,$$

where $e_\Omega \in L^\infty$ is such that $e_\Omega(\omega) = 1$ for all $\omega \in \Omega$.

The relevance of these properties is widely discussed by Artzner et al. (1999). Furthermore, the following property of a risk measure ρ has been defined by, e.g., Delbaen (2000):

- *Comonotonic Additivity*: For all $X, Y \in L^\infty$ such that X and Y are comonotone, we have that

$$\rho(X+Y) = \rho(X) + \rho(Y).$$

Random variables X and Y are comonotone if the inequality $[X(\omega_1) - X(\omega_2)] \cdot [Y(\omega_1) - Y(\omega_2)] \geq 0$ holds almost surely for $(\omega_1, \omega_2) \in \Omega \times \Omega$ with respect to the product measure $\mathbb{P} \times \mathbb{P}$ (e.g., Delbaen, 2000). If a risk measure satisfies *Comonotonic Additivity*, it means that if stochastic variables are "perfectly" dependent, there is no benefit from pooling.

Artzner et al. (1999) and Delbaen (2000) show that a risk measure ρ is coherent if and only if there exists a set of probability measures $Q \subset \mathcal{P}(\Omega, \mathcal{F})$ such that[6]

$$\rho(X) = \sup \left\{ E_{\mathbb{Q}}[X] : \mathbb{Q} \in Q \right\}, \qquad \text{for all } X \in L^\infty. \tag{1}$$

The set Q need not be unique. We will refer to a set Q that satisfies (1) as a generating probability measure set of ρ. Moreover, Delbaen (2000) shows that for every coherent risk measure ρ satisfying *Comonotonic Additivity*, there is a submodular function $\upsilon^\rho : \mathcal{F} \to \mathbb{R}_+$ with $\upsilon^\rho(\emptyset) = 0$ and $\upsilon^\rho(\Omega) = 1$ such that the following set Q is generating ρ:[7][8]

$$Q = \left\{ \mathbb{Q} \in \mathcal{P}(\Omega, \mathcal{F}) : \mathbb{Q}(A) \geq \upsilon^\rho(A) \text{ for all } A \in \mathcal{F} \right\}. \tag{2}$$

In all examples of this paper, we focus on a special class of coherent risk measures satisfying *Comonotonic Additivity*. This class is the class distortion risk measures (Wang, 1995) with a distortion function g^ρ.[9] It can be shown that, subject to a technical condition, any coherent risk measure satisfying *Comonotonic Additivity* can be represented by a distortion risk measure (Wang, Panjer and Young, 1997).

[6] This result is shown by Artzner et al. (1999) in case of a finite state space and generalized by Delbaen (2000) to stochastic variables on L^∞.

[7] This result is deduced by Delbaen (2000) from earlier results of Denneberg (1994) and Schmeidler (1986).

[8] A function $\upsilon : \mathcal{F} \to \mathbb{R}$ is submodular if $\upsilon(S \cup T) + \upsilon(S \cap T) \geq \upsilon(S) + \upsilon(T)$ for all $S, T \in \mathcal{F}$.

[9] Distortion risk measures are given by $\rho(X) = \int_0^\infty g^\rho(\mathbb{P}(X > x))dx + \int_{-\infty}^0 (g^\rho(\mathbb{P}(X > x)) - 1)dx$ for all $X \in L^\infty$, where g^ρ is a continuous, concave and increasing function such that $g^\rho(0) = 0$ and $g^\rho(1) = 1$. Here, convergence is guaranteed by boundedness of X.

Example 2.1 *For distortion risk measures, Denneberg (1994) shows that a function υ^ρ satisfying (2) is given by*

$$\upsilon^\rho(A) = 1 - g^\rho(1 - \mathbb{P}(A)), \qquad \text{for all } A \in \mathcal{F}, \tag{3}$$

where g^ρ is the distortion function.[10] *So, a function υ^ρ as in (2) has a known functional form that is only dependent on the function g^ρ and the probability space. Substituting (3) in (2) yields the following generating probability measure set of ρ:*

$$Q = \{\mathbb{Q} \in \mathcal{P}(\Omega, \mathcal{F}) : \mathbb{Q}(A) \leq g^\rho(\mathbb{P}(A)) \text{ for all } A \in \mathcal{F}\}. \tag{4}$$

Next, we discuss a well-known coherent risk measure. We use this measure in all examples in this paper. The risk measure Expected Shortfall (e.g., Acerbi and Tasche, 2002) is defined as follows.[11] *Let $\alpha \in (0,1)$. The $(1-\alpha)$-quantile is defined by*

$$q_{1-\alpha}(X) = \sup\{x \in \mathrm{I\!R} : \mathbb{P}(\{\omega \in \Omega : X(\omega) \geq x\}) > \alpha\}, \qquad \text{for all } X \in L^\infty. \tag{5}$$

Then, the risk measure Expected Shortfall with significance level $\alpha \in (0,1)$ is defined as

$$\rho_\alpha^{ES}(X) = \alpha^{-1}(E_{\mathbb{P}}\left[X \cdot \mathbb{1}_{X \geq q_{1-\alpha}(X)}\right] - q_{1-\alpha}(X) \cdot (\mathbb{P}[X \geq q_{1-\alpha}(X)] - \alpha)), \qquad \text{for all } X \in L^\infty. \tag{6}$$

Note that if X is continuously distributed, we have $\rho_\alpha^{ES}(X) = E[X : X \geq q_{1-\alpha}(X)]$. Tasche (2002) shows that this risk measure ρ is coherent and, moreover, that it satisfies Comonotonic Additivity.

Dhaene et al. (2006) show that Expected Shortfall with significance level $\alpha \in (0,1)$ is a distortion risk measure and its distortion function is given by

$$g^{\rho_\alpha^{ES}}(x) = \min\left\{\frac{x}{\alpha}, 1\right\}. \tag{7}$$

If the state space Ω is finite and the σ-algebra $\mathcal{F}$ equals its power set, we can replace the event $A \in \mathcal{F}$ in expression (4) by state $\omega \in \Omega$. This holds since $g(x)$ is linear for $x \leq \alpha$ and $\mathbb{Q}(A) \leq 1$ for all $\mathbb{Q} \in \mathcal{P}(\Omega, \mathcal{F})$ and for all $A \in \mathcal{F}$. Hence, in this case, the generating probability measure set of ρ_α^{ES} from (4) is given by

$$Q = \left\{\mathbb{Q} \in \mathcal{P}(\Omega, \mathcal{F}) : \mathbb{Q}(\{\omega\}) \leq \frac{\mathbb{P}(\{\omega\})}{\alpha} \text{ for all } \omega \in \Omega\right\}. \tag{8}$$

In Example 2.4, we will discuss this probability measure set in more detail. $\triangledown$

Next, we introduce a special class of risk measures. In this paper, we consider risk measures that are *finitely generated*, i.e., the risk measure has a finite generating probability measure set Q.

Definition 2.2 *A coherent risk measure $\rho : L^\infty(\Omega, \mathcal{F}, \mathbb{P}) \to \mathrm{I\!R}$ is finitely generated if there exists a finite generating probability measure set $Q \subset \mathcal{P}(\Omega, \mathcal{F})$, i.e.,*

$$\rho(X) = \max\{E_{\mathbb{Q}}[X] : \mathbb{Q} \in Q\}, \qquad \text{for all } X \in L^\infty. \tag{9}$$

Again, we note that this set Q need not be uniquely determined.

The condition in Definition 2.2 on coherent risk measures seems quite restrictive. However, we show that all coherent risk measures satisfying *Comonotonic Additivity* satisfy this property in case the state space is finite.

Proposition 2.3 *If the state space Ω is finite and the risk measure ρ is coherent and satisfies Comonotonic Additivity, then ρ is finitely generated.*

[10] This result is shown by Denneberg (1994) for Choquet integrals. The distortion risk measure is an example of a Choquet integral.

[11] The risk measure Expected Shortfall is also referred by other authors as Worst Conditional Expectation (e.g., Artzner et al., 1999), Conditional VaR or Tail VaR (e.g., Dhaene et al., 2006).

Proof Let Q be the generating probability measure set of ρ that is defined in (2), i.e.,

$$Q = \{\mathbb{Q} \in \mathcal{P}(\Omega, \mathcal{F}) : \mathbb{Q}(A) \geq \upsilon^\rho(A) \text{ for all } A \in \mathcal{F}\},$$

where $\upsilon^\rho : \mathcal{F} \to \mathbb{R}_+$ is submodular, $\upsilon^\rho(\emptyset) = 0$ and $\upsilon^\rho(\Omega) = 1$. Note that as the state space Ω is finite, the σ-algebra $\mathcal{F}$ is finite as well. Because $\mathcal{F}$ is finite, Q is defined via a finite number of linear inequalities on $[0,1]^\Omega$. So, Q is a convex polytope. Let $\widetilde{Q}$ be the finite collection of extreme points of this convex polytope. Because $\mathbb{Q} \to E_\mathbb{Q}[X]$ is a linear map on Q for every $X \in L^\infty$, (1) is a linear programming problem and, therefore, we have

$$\rho(X) = \sup\{E_\mathbb{Q}[X] : \mathbb{Q} \in Q\} = \max\left\{E_\mathbb{Q}[X] : \mathbb{Q} \in \widetilde{Q}\right\}, \qquad \text{for all } X \in L^\infty.$$

Hence, $\rho(X)$ equals the maximum of all expectations of X under the probability measures in $\widetilde{Q}$. Hence, $\widetilde{Q}$ is a generating probability measure set. This concludes the proof. □

A finitely generated risk measure, however, does not need to satisfy *Comonotonic Additivity*. As a main example, Expected Shortfall belongs to the class of finitely generated risk measures if Ω is finite.

In the next example, we provide an explicit expression for a specific finite generating probability measure set corresponding to Proposition 2.3.

Example 2.4 *Let the state space Ω be finite, the σ-algebra $\mathcal{F}$ equal its power set and the risk measure ρ be a coherent risk measure satisfying Comonotonic Additivity. Recall the generating probability measure set in (2). Let σ be an order on Ω, i.e., $\sigma : \{1, \ldots, |\Omega|\} \to \Omega$ is the bijective function that corresponds with a permutation on the state space. The state at position $j \in \{1, \ldots, |\Omega|\}$ in the order σ is denoted by $\sigma(j) \in \Omega$ and the set of all orders on the state space Ω is denoted by $\Pi(\Omega)$. We define for every order $\sigma \in \Pi(\Omega)$ the following stochastic variable:*

$$m^\sigma(\sigma(j)) = \upsilon^\rho\left(\bigcup_{k=1}^{j}\{\sigma(k)\}\right) - \upsilon^\rho\left(\bigcup_{k=1}^{j-1}\{\sigma(k)\}\right), \qquad \text{for all } j \in \{1, \ldots, |\Omega|\}. \tag{10}$$

The stochastic variable m^σ is a probability measure on $(\Omega, \mathcal{F})$ due to submodularity and non-negativity of υ^ρ, $\upsilon^\rho(\emptyset) = 0$ and $\upsilon^\rho(\Omega) = 1$. For distortion risk measures, we can simplify (10) to

$$m^\sigma(\sigma(j)) = g^\rho\left(\sum_{k=j}^{|\Omega|}\mathbb{P}(\{\sigma(k)\})\right) - g^\rho\left(\sum_{k=j+1}^{|\Omega|}\mathbb{P}(\{\sigma(k)\})\right), \qquad \text{for all } \sigma \in \Pi(\Omega) \text{ and } j \in \{1, \ldots, |\Omega|\}. \tag{11}$$

Then, the following set is a generating probability measure set of ρ:

$$Q = \{m^\sigma : \sigma \in \Pi(\Omega)\}. \tag{12}$$

This result follows almost directly from the proof of Proposition 2.3.[12]

Next, we provide an example of the construction of (12). Let $\Omega = \{\omega_1, \omega_2, \omega_3\}$. If the state space is finite, we write probability measures as row vectors. We let the physical probability measure given by $\mathbb{P} = \left(\frac{1}{20}, \frac{9}{20}, \frac{1}{2}\right)$ and $\rho = \rho^{ES}_{0.1}$. Using (11) with $g^{\rho^{ES}_{0.1}}(x) = \min\left\{\frac{x}{0.1}, 1\right\}$ for all $x \in [0,1]$, we obtain the following outcomes of $(m^\sigma(\omega_1), m^\sigma(\omega_2), m^\sigma(\omega_3))$ for all $\sigma \in \Pi(\Omega)$:

σ	ω_1	ω_2	ω_3
$\omega_1\,\omega_2\,\omega_3$	0	0	1
$\omega_1\,\omega_3\,\omega_2$	0	1	0
$\omega_2\,\omega_1\,\omega_3$	0	0	1
$\omega_2\,\omega_3\,\omega_1$	$\frac{1}{2}$	0	$\frac{1}{2}$
$\omega_3\,\omega_2\,\omega_1$	$\frac{1}{2}$	$\frac{1}{2}$	0
$\omega_3\,\omega_1\,\omega_2$	0	1	0

[12]Suppose (Ω, υ^ρ) is a Transferable Utility game, where Ω is the corresponding "player" set. In game-theoretical terms, the representation (2) of a finite generating probability measure set coincides with the core of the game (Ω, υ^ρ). Then, submodularity of the function υ^ρ is equivalent with convexity of the corresponding game (Shapley, 1971). Shapley (1971) shows that the core of convex games coincides with the convex hull of the marginal vectors. A marginal vector corresponds with a vector m^σ. All extreme points of the core of (Ω, υ^ρ) are in (12).

Hence, according to (12), a finite generating probability measure set of $\rho_{0.1}^{ES}$ is given by

$$Q = \{\mathbb{Q}_1, \mathbb{Q}_2, \mathbb{Q}_3, \mathbb{Q}_4\}, \tag{13}$$

where $\mathbb{Q}_1 = \left(\frac{1}{2}, \frac{1}{2}, 0\right)$, $\mathbb{Q}_2 = (0, 0, 1)$, $\mathbb{Q}_3 = (0, 1, 0)$, *and* $\mathbb{Q}_4 = \left(\frac{1}{2}, 0, \frac{1}{2}\right)$. *Recall that the elements of Q are the extreme points of the set in (8).* $\bigtriangledown$

Throughout the sequel of this paper, we fix for a given finitely generated risk measure ρ a finite generating probability measure set, which we denote by $Q(\rho)$. This assumption is without loss of generality.

2.2 Risk capital allocation problems

In this section, we discuss risk capital allocation problems as in, e.g., Denault (2001). Consider a firm, for example a pension fund or an insurance company. This firm consists of multiple divisions that face risk. The risk of a division is summarized by a stochastic loss variable at a common future time. The problem is to allocate the total risk among all divisions.

The finite set of all divisions within a firm is denoted by N. Throughout this paper, we fix N and the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We denote for each division $i \in N$ the stochastic loss as $X_i \in L^\infty$. The total loss of the firm is given by $\sum_{i \in N} X_i$. We assume that the risk capital of the firm is measured using a coherent risk measure ρ. In the following definition, we define the risk capital allocation problem.

Definition 2.5 *A risk capital allocation problem is a tuple $((X_i)_{i\in N}, \rho)$, where $X_i \in L^\infty$ for all $i \in N$ and ρ is a finitely generated risk measure. The class of all risk capital allocation problems is denoted by $\mathcal{R}$.*

In the following definition, we define the concept of *risk capital allocations* and *risk capital allocation rules.*

Definition 2.6 *A vector $a \in \mathbb{R}^N$ is a risk capital allocation for $R = ((X_i)_{i\in N}, \rho) \in \mathcal{R}$ if and only if $\sum_{i\in N} a_i = \rho\left(\sum_{i\in N} X_i\right)$. Let $\widetilde{\mathcal{R}} \subseteq \mathcal{R}$ be a (sub)domain of risk capital allocation problems. We define a risk capital allocation rule as a function $K : \widetilde{\mathcal{R}} \to \mathbb{R}^N$ that assigns to every risk capital allocation problem $R \in \widetilde{\mathcal{R}}$ a unique risk capital allocation $K(R)$. So, we get*

$$\sum_{i\in N} K_i(R) = \rho\left(\sum_{i\in N} X_i\right), \qquad \textit{for all } R \in \widetilde{\mathcal{R}}.$$

The *Sub-additivity* property of coherent risk measures implies that there can be benefits from pooling risks. Specifically, it implies that

$$\rho\left(\sum_{i\in N} X_i\right) \leq \sum_{i\in N} \rho(X_i). \tag{14}$$

This property implies that allocating risk capital among divisions is generally non-trivial. The aim is to allocate the gains from pooling risk in a fair way.

Based on Denault (2001), we define the following properties of a risk capital allocation rule $K : \widetilde{\mathcal{R}} \to \mathbb{R}^N$:

- *Translation Invariance*: For all $R = ((X_i)_{i\in N}, \rho) \in \widetilde{\mathcal{R}}$, it holds that if $\hat{R} = ((\hat{X}_i)_{i\in N}, \rho) \in \widetilde{\mathcal{R}}$ where $(\hat{X}_i)_{i\in N} = (X_j + c \cdot e_\Omega, X_{-j})$ for some $c \in \mathbb{R}$ and $j \in N$, then

  $$K(\hat{R}) = K(R) + c \cdot e_j,$$

  where e_j is the j-th unit vector in $\mathbb{R}^N$.

- *Scale Invariance*: For all $R = ((X_i)_{i\in N}, \rho) \in \widetilde{\mathcal{R}}$, it holds that if $\hat{R} = ((\hat{X}_i)_{i\in N}, \rho) \in \widetilde{\mathcal{R}}$ where $(\hat{X}_i)_{i\in N} = (c \cdot X_i)_{i\in N}$ for some $c > 0$, then

  $$K(\hat{R}) = c \cdot K(R).$$

- *Monotonicity*: For all $R \in \widetilde{\mathcal{R}}$ where ρ is non-decreasing in the sense that $\rho\left(\sum_{i\in N} \lambda_i X_i\right) \leq \rho\left(\sum_{i\in N} \lambda_i^* X_i\right)$ whenever $\lambda, \lambda^* \in [0,1]^N$ and $\lambda \leq \lambda^*$, we have

  $$K(R) \geq 0.$$

- *Fuzzy Core Selection*: For all $R \in \widetilde{R}$, we have

  $$K(R) \in FCore(R), \qquad \text{for all } R \in \widetilde{\mathcal{R}},$$

  where $FCore(R)$ denotes the fuzzy core (Aubin, 1979), which is defined as:[13]

  $$FCore(R) = \left\{ a \in \mathbb{R}^N : \sum_{i\in N} \lambda_i a_i \leq \rho\left(\sum_{i\in N} \lambda_i X_i\right) \text{ for all } \lambda \in [0,1]^N, \sum_{i\in N} a_i = \rho\left(\sum_{i\in N} X_i\right)\right\}. \tag{15}$$

In particular, the property *Fuzzy Core Selection* is widely discussed in the game-theoretic literature on risk capital allocation problems (see, e.g., Denault, 2001; Tsanakas and Barnett, 2003). For an allocation in the fuzzy core, no portfolio of fractional risks would have a lower stand-alone risk capital than the corresponding risk capital allocation. This property ensures a stable allocation.

3 The Shapley value and the Aumann-Shapley value

In this section, we discuss cooperative game-theoretic solution concepts for risk capital allocation problems. There is one particular solution concept that has received considerable attention in cooperative game theory, namely the Shapley value (Shapley, 1953). The following definition defines the Shapley value for risk capital allocation problems.

Definition 3.1 *The Shapley value for risk capital allocation problems, denoted by $S : \mathcal{R} \to \mathbb{R}^N$, is given by*

$$S_i(R) = \sum_{S\subseteq N\setminus\{i\}} w(|S|) \cdot \left(\rho\left(\sum_{j\in S\cup\{i\}} X_j\right) - \rho\left(\sum_{j\in S} X_j\right)\right),$$

for all $R \in \mathcal{R}$ and $i \in N$, where $|S|$ denotes the number of divisions in $S \subseteq N$ and

$$w(|S|) = \frac{|S|! \cdot (|N| - |S| - 1)!}{|N|!}.$$

The Shapley value is originally defined for Transferable Utility games and is here applied to the atomic risk capital cost game of Denault (2001).[14] The Shapley value can be interpreted as the average of all marginal vectors. Given an ordering of divisions, a marginal vector is created by assigning to every division its marginal contribution if they enter the coalition one-by-one according to the order. The weight $w(|S|)$ assigned to a marginal contribution of division i to a coalition S represents the "probability" that in a uniformly random ordering of divisions, all divisions in S are on the first positions and thereafter is division i.

It is easy to show that the Shapley value is a risk capital allocation rule. Denault (2001) shows that the Shapley value does not satisfy the stability criterium that, in game-theoretical terms, means that the Shapley value may not yield a core element. This is seen as a major drawback of the Shapley value for risk capital allocation problems.

Aumann and Shapley (1974) developed an allocation rule for general cost functions. This is given in the following definition.

Definition 3.2 *The Aumann-Shapley value of a function $r : [0,1]^N \to \mathbb{R}$ is given by $(a_i)_{i\in N}$ such that*

$$a_i = \int_0^1 \frac{\partial r}{\partial \lambda_i}(\gamma \cdot e_N) d\gamma, \qquad \text{for all } i \in N, \tag{16}$$

whenever these integrals exist, and where e_N is the unit vector in $\mathbb{R}^N$.

[13]Aubin (1979) shows that the fuzzy core is non-empty, convex and compact since ρ satisfies *Sub-additivity* and *Positive Homogeneity*.
[14]Denault (2001) defines the atomic risk capital game (N, c) as $c(S) = \rho(\sum_{i\in S} X_i)$ for all $S \subseteq N$.

The Aumann-Shapley value can be interpreted as the average of the marginal changes of the risk capital function r if the participation of all divisions increases simultaneously.

Denault (2001) applies the Aumann-Shapley value to risk capital allocation problems. The *risk capital function* $r : [0,1]^N \to \mathbb{R}$ of a risk capital allocation problem $R = ((X_i)_{i\in N}, \rho) \in \mathcal{R}$ is defined as

$$r(\lambda) = \rho\left(\sum_{i\in N} \lambda_i \cdot X_i\right), \qquad \text{for all } \lambda \in [0,1]^N. \tag{17}$$

The interpretation is as follows. Here, divisions are allowed to participate fractionally. The participation level for division $i \in N$, denoted by $\lambda_i \in [0,1]$, can be seen as the fractional involvement of division i in a coalition. Here, the maximal participation level of every division is normalized to one.[15] Moreover, we define a participation profile $\lambda \in [0,1]^N$ as a collection of participation levels of all divisions in N. Note that the risk capital of the firm is given by $r(e_N) = \rho(\sum_{i\in N} X_i)$.

Since the risk measure ρ satisfies *Positive Homogeneity*, we can simplify expression (16) as in the following corollary of, e.g., Denault (2001).

Corollary 3.3 *The Aumann-Shapley value for risk capital allocation problems, denoted by $AS : \mathcal{R}' \to \mathbb{R}^N$, is given by*

$$AS_i(R) = \frac{\partial r}{\partial \lambda_i}(e_N), \qquad \textit{for all } i \in N, \tag{18}$$

where the risk capital function r is as in (17) and $\mathcal{R}' \subset \mathcal{R}$ is defined as the set of all risk capital allocation problems for which r is partially differentiable at $\lambda = e_N$.

The Aumann-Shapley value for risk capital allocation problems is the gradient of the risk capital function evaluated in the vector of full participation. As we focus on risk capital allocation problems in this paper, we will continue by referring to AS as the Aumann-Shapley value.

We discuss the main game-theoretic argument why the Aumann-Shapley value is widely supported as allocation rule for risk capital. This property involves the fuzzy core that is defined in (15). Allocations in the fuzzy core are derived on the premise that every fuzzy portfolio should be allocated less than its stand-alone risk capital, i.e., no fuzzy coalition λ has an incentive to split off from the firm. The following theorem about the relationship between the fuzzy core and the Aumann-Shapley value is due to Denault, which is based on an earlier result of Aubin (1979).

Theorem 3.4 (Denault, 2001, Theorem 7, page 20) *For all $R \in \mathcal{R}'$, the fuzzy core $FCore(R)$ consists of only one element. This single-valued fuzzy core element is the Aumann-Shapley value.*

The worst-case probability measures for the firm is defined as follows:

$$Q^*(\rho) = \left\{\mathbb{Q} \in Q(\rho) : r(e_N) = E_{\mathbb{Q}}\left[\sum_{i\in N} X_i\right]\right\}. \tag{19}$$

The following theorem provides an expression of the fuzzy core. This is based on Aubin (1979, Proposition 4, page 343).

Theorem 3.5 *Let $R \in \mathcal{R}$. Then, we have*

$$FCore(R) = conv\{(E_{\mathbb{Q}}[X_i])_{i\in N} : \mathbb{Q} \in Q^*(\rho)\},$$

where conv denotes the convex hull operator and $Q^*(\rho)$ is defined in (19).

[15] This normalization is contrary to the approach of Denault (2001), but it is without loss of generality.

Next, we can reformulate the Aumann-Shapley value from (18). It follows from Theorem 3.4 and Theorem 3.5 that the Aumann-Shapley value for division $i \in N$, if well-defined, is given by the expectation of X_i under the worst-case probability measure $\mathbb{Q} \in Q^*(\rho)$, i.e.,

$$AS_i(R) = E_{\mathbb{Q}}[X_i], \qquad \text{for all } R \in \mathcal{R}', i \in N \text{ and } \mathbb{Q} \in Q^*(\rho). \tag{20}$$

If there exists a division i and $\mathbb{Q}, \widetilde{\mathbb{Q}} \in Q^*(\rho)$ such that $E_{\mathbb{Q}}[X_i] \neq E_{\widetilde{\mathbb{Q}}}[X_i]$, then the Aumann-Shapley value does not exist. Moreover, the following result is based on a characterization of the Aumann-Shapley value of Denault (2001).

Theorem 3.6 (Denault, 2001, Corollary 1, page 20) *The Aumann-Shapley value satisfies the properties Translation Invariance, Scale Invariance and Monotonicity on $\mathcal{R}'$.*

Next, we provide an example about the construction of the Aumann-Shapley value.

Example 3.7 *Recall the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ from Example 2.4. Let $N = \{1, 2\}$. If the state space is finite, we write stochastic variables as column vectors. Let the risk capital allocation problem $R = ((X_i)_{i \in N}, \rho)$ be given by*

$$X_1 = \begin{bmatrix} 0 \\ 2 \\ 4 \end{bmatrix}, X_2 = \begin{bmatrix} 6 \\ 2 \\ -2 \end{bmatrix}, \text{ and, again, } \rho = \rho_\alpha^{ES}.$$

The corresponding risk capital function r, which is defined in (17), is given by

$$r(\lambda_1, \lambda_2) = \begin{cases} \lambda_1 + 4\lambda_2 & \text{if } 2\lambda_2 \geq \lambda_1, \\ 4\lambda_1 - 2\lambda_2 & \text{otherwise}, \end{cases} \tag{21}$$

for all $\lambda \in [0, 1]^N$. This fuzzy risk capital function is displayed in Figure 1. Note that the participation profiles λ

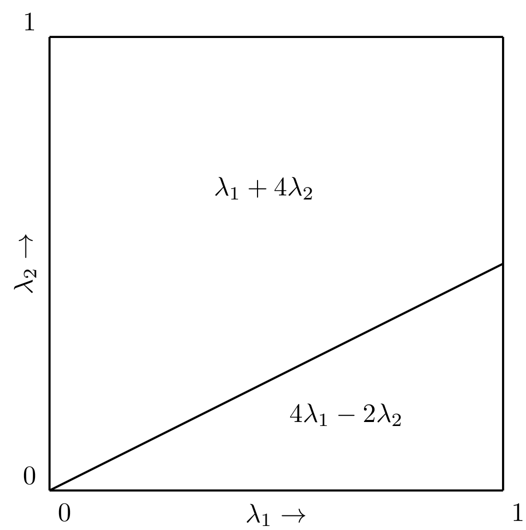


Figure 1: Plot of $r(\lambda_1, \lambda_2)$ corresponding to (21).

where the risk capital function r is not partially differentiable are located on a straight line through the origin. It can be verified that $\mathbb{Q} = \left(\frac{1}{2}, \frac{1}{2}, 0\right)$ is the unique probability measure in $Q^*(\rho_{0.1}^{ES})$ from (19). Hence, using (20), the Aumann-Shapley value is given by

$$AS(R) = (E_{\mathbb{Q}}[X_1], E_{\mathbb{Q}}[X_2]) = (1, 4).$$

$\bigtriangledown$

The main drawback of the Aumann-Shapley value is that it requires partial differentiability of the risk capital function. In Section 6, we extend the Aumann-Shapley value such that it is always well-defined.

4 Fuzzy risk capital allocation functions

In this section, we analyze the structure of the risk capital function r, which is defined in (17). We show that it is piecewise linear and almost everywhere partially differentiable. We also introduce some notation that we need in Section 6, where we define an allocation rule.

Definition 4.1 *Let $R = ((X_i)_{i\in N}, \rho) \in \mathcal{R}$. Then, we define for all $\mathbb{Q} \in Q(\rho)$ the linear function $f_{\mathbb{Q}} : \mathbb{R}^N \to \mathbb{R}$ by*

$$f_{\mathbb{Q}}(\lambda) = \sum_{i\in N} \lambda_i \cdot E_{\mathbb{Q}}[X_i], \qquad \text{for all } \lambda \in \mathbb{R}^N.$$

Note that it is possible that different measures $\mathbb{Q} \in Q(\rho)$ yield the same function $f_{\mathbb{Q}}$.

Proposition 4.2 *For all $R \in \mathcal{R}$, the risk capital function r is piecewise linear on $[0,1]^N$.*

Proof For all $R \in \mathcal{R}$, we have

$$\begin{aligned} r(\lambda) &= \max\left\{ E_{\mathbb{Q}}\left[\sum_{i\in N} \lambda_i \cdot X_i\right] : \mathbb{Q} \in Q(\rho)\right\} \\ &= \max\left\{ \sum_{i\in N} \lambda_i \cdot E_{\mathbb{Q}}\left[X_i\right] : \mathbb{Q} \in Q(\rho)\right\} \\ &= \max\{f_{\mathbb{Q}}(\lambda) : \mathbb{Q} \in Q(\rho)\}, \end{aligned} \tag{22}$$

for all $\lambda \in [0,1]^N$. This concludes the proof. $\square$

From (22) follows that for every $\lambda \in [0,1]^N$, there exists at least one $\mathbb{Q} \in Q(\rho)$ such that $r(\lambda) = f_{\mathbb{Q}}(\lambda)$.

Next, we define the set of participation profiles corresponding to a probability measure in $Q(\rho)$ where this measure is the worst case probability measure.

Definition 4.3 *Let $R = ((X_i)_{i\in N}, \rho) \in \mathcal{R}$. Then, we define for all $\mathbb{Q} \in Q(\rho)$ the participation profile set $A_{\mathbb{Q}} \subseteq [0,1]^N$ as*

$$A_{\mathbb{Q}} = \left\{\lambda \in [0,1]^N : r(\lambda) = f_{\mathbb{Q}}(\lambda)\right\}.$$

Moreover, we define and order $K_1, \dots, K_p$, $p^* \leq p$ and $\mathbb{Q}_1, \dots, \mathbb{Q}_p \in Q(\rho)$ such that:

- *$K_1, \dots, K_p$ is an exhaustive list of the elements of $\{A_{\mathbb{Q}} : \mathbb{Q} \in Q(\rho)\}$ without repetitions;*
- *$K_m = A_{\mathbb{Q}_m}$ for all $m \in \{1, \dots, p\}$;*
- *$e_N \in K_m$ for $m \in \{1, \dots, p^*\}$ and $e_N \notin K_m$ otherwise.*

It is straightforward to show that K_m is a closed and convex polytope (in fact, a pointed cone). Define $e_\emptyset$ as the zero vector in $\mathbb{R}^N$. The participation profile $\lambda = e_\emptyset$ is an element of K_m for all $m \in \{1, \dots, p\}$. Moreover, we have

$$\bigcup_{m=1}^{p} K_m = [0,1]^N, \tag{23}$$

and $p \leq |Q(\rho)|$.

Example 4.4 *In this example, we illustrate Definition 4.3. Recall the probability space $(\Omega, \mathcal{F}, \mathbb{P})$,the risk capital allocation problem from Example 3.7 and $\mathbb{Q}_1, \dots, \mathbb{Q}_4$ from Example 2.4. The ordering of $\mathbb{Q}_1, \dots, \mathbb{Q}_4$ corresponds with Definition 4.3. Based on this definition, we get $K_1 = A_{\mathbb{Q}_1} = \left\{\lambda \in [0,1]^N : \lambda_1 \leq 2\lambda_2\right\}$, $K_2 = A_{\mathbb{Q}_2} = \left\{\lambda \in [0,1]^N : \lambda_1 \geq 2\lambda_2\right\}$ and $K_3 = A_{\mathbb{Q}_3} = A_{\mathbb{Q}_4} = \left\{\lambda \in [0,1]^N : \lambda_1 = 2\lambda_2\right\}$. Note that $f_{\mathbb{Q}_3} = f_{\mathbb{Q}_4}$ and, so, it is without loss of generality to drop $\mathbb{Q}_4$.* $\triangledown$

Next, we focus on partial differentiability of the risk capital function r. Partial differentiability is a key issue for existence of the Aumann-Shapley value (see Corollary 3.3) and, moreover, it is key in the risk capital allocation rule that we define in Section 6.

Definition 4.5 *Let $R \in \mathcal{R}$. Then, the set $L(R)$ is given by*

$$L(R) = \left\{\lambda \in [0,1]^N : \textit{there exists a unique } m \in \{1,\dots,p\} \textit{ such that } \lambda \in K_m\right\}.$$

The set $L(R)$ is open on $[0,1]^N$ and, therefore, $L(R)$ is a the set of participation profiles where risk capital function r is locally linear. So, for all $\lambda \in L(R) \cap K_m$, there exists a neighborhood $U \subset [0,1]^N$ of λ such that $r(\hat{\lambda}) = f_{\mathbb{Q}_m}(\hat{\lambda})$ for all $\hat{\lambda} \in U$ and, so,

$$\frac{\partial r}{\partial \lambda_i}(\lambda) = E_{\mathbb{Q}_m}[X_i], \qquad \text{for all } i \in N. \tag{24}$$

Hence, $L(R)$ is a set of participation profiles in $[0,1]^N$ on which the risk capital function r is partially differentiable.[16] Note that $p^* = 1$ if and only if the function r is partially differentiable in $\lambda = e_N$.

Lemma 4.6 *For all $R \in \mathcal{R}$, the collection of profiles where the risk capital function r is not partially differentiable is a subset of a collection of a finite number of hyperplanes passing trough $\lambda = e_\emptyset$.*

Proof We obtain for all $\ell, m \in \{1,\dots,p\}$ that

$$\begin{aligned} K_\ell \cap K_m &= \left\{\lambda \in [0,1]^N : r(\lambda) = f_{\mathbb{Q}_\ell}(\lambda) = f_{\mathbb{Q}_m}(\lambda)\right\} \\ &\subseteq \left\{\lambda \in [0,1]^N : f_{\mathbb{Q}_\ell}(\lambda) = f_{\mathbb{Q}_m}(\lambda)\right\} \\ &= \left\{\lambda \in [0,1]^N : \sum_{i \in N} \lambda_i \cdot \left(E_{\mathbb{Q}_\ell}[X_i] - E_{\mathbb{Q}_m}[X_i]\right) = 0\right\}. \end{aligned} \tag{25}$$

If $E_{\mathbb{Q}_\ell}[X_i] = E_{\mathbb{Q}_m}[X_i]$ for all $i \in N$, we have $K_\ell = K_m$ which implies $\ell = m$. So, the set $K_\ell \cap K_m$ is a (possibly empty) subset of a hyperplane passing through $\lambda = e_\emptyset$ for all $\ell, m \in \{1,\dots,p\}$ such that $\ell \neq m$. We have by construction that

$$[0,1]^N \backslash L(R) = \bigcup_{\ell,m \in \{1,\dots,p\} : \ell \neq m} K_\ell \cap K_m, \qquad \text{for all } R \in \mathcal{R}. \tag{26}$$

From this follows that the collection of profiles where the risk capital function r is not partially differentiable is a subset of the collection of a finite number of hyperplanes passing trough $\lambda = e_\emptyset$. □

From this lemma, we get that the set $L(R)$ has a full measure on $[0,1]^N$. So, we obtain immediately the following corollary.

Corollary 4.7 *For all $R \in \mathcal{R}$, the risk capital function r is almost everywhere partially differentiable.*

In the sequel of this paper, we use the risk capital function r for defining an allocation rule in Section 6. We will need the results in Lemma 4.6 and Corollary 4.7 in Subsection 6.2.

5 Path based allocation rules

In this section, we discuss *path based allocation rules* as introduced by Wang (1999). We construct an allocation based on the idea of the marginal vectors of the Shapley value. We extend this idea to a problem where divisions

[16] The set $L(R)$ is not necessary the set of all participation profiles where r is partially differentiable as if $\lambda_i = 0$ for some division i, there may be multiple m such that $\lambda \in K_m$ even though r may be partially differentiable.

can participate fractionally via a finite set of participation levels. We first describe this allocation rule informally and, thereafter, we provide a formal definition.

Let $n \in \mathbb{N}$ and define the grid on $[0,1]^N$ with grid size $\frac{1}{n}$ by

$$G^n = \left\{0, \frac{1}{n}, \frac{2}{n}, \ldots, 1\right\}^N. \tag{27}$$

The starting point on grid G^n is the participation profile $e_\emptyset$ in which the participation level of each division is zero. In the first step the participation level of some division i is increased by $\frac{1}{n}$ and the corresponding difference in risk capital, $r((1/n) \cdot e_i) - r(e_\emptyset)$, is allocated to division i. In the second step of the path again the participation level of some division (not necessarily the same as the one in the first step) is increased by $\frac{1}{n}$ and the risk change is allocated to this division. Proceeding in this way, we will end up after $|N|n$ steps in e_N and risk capital $r(e_N)$ has been allocated to the divisions by then.

Formally, we define a path in the following way.

Definition 5.1 *Let $n \in \mathbb{N}$. A path on the grid G^n is a map $P : \{0, 1, 2, \ldots, |N|n\} \to G^n$ satisfying:*

1. $P(0) = e_\emptyset$ *and* $P(|N|n) = e_N$;
2. *for every $k \in \{0, \ldots, |N|n - 1\}$ there exists a unique $i \in N$ such that*

$$P(k+1) - P(k) = \frac{1}{n} \cdot e_i. \tag{28}$$

This unique division i will be denoted as $i(P, k)$.

An example of a path P on the grid G^n is given in Figure 2. We denote the collection of all paths over the grid G^n by $\mathcal{P}^n$. For any path $P \in \mathcal{P}^n$, we define an allocation rule A^P on $\mathcal{R}$ by allocating the risk changes along the path

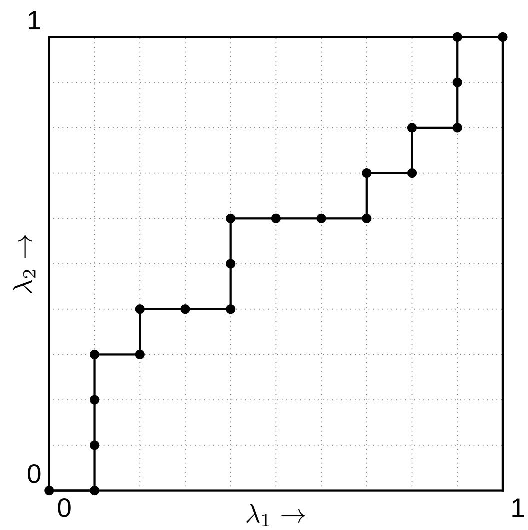


Figure 2: Example of a path $P \in \mathcal{P}^n$ for $|N| = 2$ with $n = 10$. We connected succeeding elements of the path as illustration.

to the corresponding divisions. Formally, we define this allocation rule as follows.

Definition 5.2 *For a given path $P \in \mathcal{P}^n$, the rule $A^P : \mathcal{R} \to \mathbb{R}^N$ is defined by*

$$A^P(R) = \sum_{k=0}^{|N|n-1} [r(P(k+1)) - r(P(k))] \cdot e_{i(P,k)}, \qquad \textit{for all } R \in \mathcal{R},$$

where the risk capital function r is defined in (17).

Proposition 5.3 *For every $P \in \mathcal{P}^n$, A^P is an allocation rule on $\mathcal{R}$.*

Proof Let $n \in \mathbb{N}$ and $P \in \mathcal{P}^n$. Then, the result follows directly from

$$\begin{aligned}
\sum_{i\in N} A_i^P(R) &= \sum_{i\in N} \sum_{k=0}^{|N|n-1} [r(P(k+1)) - r(P(k))] \cdot \mathbb{1}_{i(P,k)=i} && (29)\\
&= \sum_{k=0}^{|N|n-1} [r(P(k+1)) - r(P(k))] \cdot \sum_{i\in N} \mathbb{1}_{i(P,k)=i} && (30)\\
&= \sum_{k=0}^{|N|n-1} [r(P(k+1)) - r(P(k))] && (31)\\
&= r(P(|N|n)) - r(P(0))\\
&= r(e_N), && (32)
\end{aligned}$$

where $\mathbb{1}_{i(P,k)=i} = 1$ if $i(P,k) = i$ and $\mathbb{1}_{i(P,k)=i} = 0$ otherwise. Here, (29) follows from Definition 5.2, (30) follows by interchanging the summations, (31) follows from the fact that there is precisely one $i \in N$ such that $i(P,k) = i$ for all $k \in \{0, \ldots, |N|n-1\}$ and (32) follows from Definition 5.1.1. This concludes the proof. $\square$

We will refer to A^P as a path based allocation rule. Next, we show some general properties of path based allocation rules.

Theorem 5.4 *For every $P \in \mathcal{P}^n$, the allocation rule A^P satisfies the properties Translation Invariance, Scale Invariance and Monotonicity on $\mathcal{R}$.*

Proof We start with showing the property *Translation Invariance*. Let $P \in \mathcal{P}^n$, $j \in N$, $R = ((X_i)_{i\in N}, \rho) \in \mathcal{R}$ and $\hat{R} = ((\hat{X}_i)_{i\in N}, \rho) \in \mathcal{R}$ such that $(\hat{X}_i)_{i\in N} = (X_j + c \cdot e_\Omega, X_{-j})$ for some $c \in \mathbb{R}$. Moreover, we let the risk capital function r (resp. $\hat{r}$), defined in (17), correspond with R (resp. $\hat{R}$). Then, we get

$$\begin{aligned}
\hat{r}(\lambda) &= \rho\left(\sum_{i\in N} \lambda_i \hat{X}_i\right)\\
&= \rho\left(\sum_{i\in N} \lambda_i \cdot X_i + c \cdot \lambda_j \cdot e_\Omega\right)\\
&= \rho\left(\sum_{i\in N} \lambda_i \cdot X_i\right) + c \cdot \lambda_j && (33)\\
&= r(\lambda) + c \cdot \lambda_j, && (34)
\end{aligned}$$

for all $\lambda \in [0,1]^N$, where (33) follows from *Translation Invariance* of ρ. We get

$$\begin{aligned}
A^P(\hat{R}) &= \sum_{k=0}^{|N|n-1} [\hat{r}(P(k+1)) - \hat{r}(P(k))] \cdot e_{i(P,k)} && (35)\\
&= \sum_{k=0}^{|N|n-1} [r(P(k+1)) + c \cdot P_j(k+1) - r(P(k)) - c \cdot P_j(k)] \cdot e_{i(P,k)} && (36)\\
&= A^P(R) + c \cdot \sum_{k=0}^{|N|n-1} [P_j(k+1) - P_j(k)] \cdot e_{i(P,k)} && (37)\\
&= A^P(R) + c \cdot \sum_{k=0}^{|N|n-1} [P_j(k+1) - P_j(k)] \cdot e_j && (38)
\end{aligned}$$

$$= A^P(R) + c \cdot [P_j(|N|n) - P_j(0)] \cdot e_j$$
$$= A^P(R) + c \cdot e_j, \tag{39}$$

where $P_j(k)$ is the j-th element of $P(k)$. Here, (35) follows from Definition 5.2, (36) follows from (118), (37) follows from Definition 5.2, (38) follows from $P_j(k+1) - P_j(k) = 0$ if $i(P,k) \neq j$ (see (28)) and (39) follows from Definition 5.1.1. This concludes the proof of *Translation Invariance*.

The proof of *Scale Invariance* is similar to the proof of *Translation Invariance*.

Next, we show *Monotonicity*. Let the risk measure ρ be non-decreasing. This implies $r(P(k+1)) - r(P(k)) \geq 0$ for all $k \in \{0, \ldots, |N|n-1\}$. Then, from Definition 5.2 it follows directly that $A^P(R) \geq 0$. This concludes the proof of *Monotonicity*. $\square$

Compare this result with Theorem 3.6. Denault (2001) shows that there exists an allocation rule satisfying the above-mentioned properties and that the Aumann-Shapley value is such a rule. Using Theorem 5.4, we extend this result for all allocation rules based on a path. Moreover, one can verify using *Positive Homogeneity* and *Sub-additivity* of ρ that a path based allocation rule always satisfies the following bounds:

$$\min\{E_{\mathbb{Q}}[X_i] : \mathbb{Q} \in Q(\rho)\} \leq A_i^P(R) \leq \rho(X_i),$$

for all $P \in \mathcal{P}^n$, $R = ((X_i)_{i \in N}, \rho) \in \mathcal{R}$ and $i \in N$. So, every path based allocation is individually rational and is always weakly more than the expectation of X_i under the best-case probability measure.

We can approximate expression (16) of the Aumann-Shapley value (if existent) using a very small grid and a path close to the diagonal. Note that the Aumann-Shapley value is not defined if the risk capital function r is not partially differentiable along the diagonal. As a solution, we propose a generalization based on paths that is not prone to this problem in the next section.

6 The Weighted-Aumann Shapley value

6.1 A sequence of discrete rules

In this section we introduce a generalization of the well-known Aumann-Shapley value. In line with the Shapley value (1953), we define an allocation rule based on the average of all path based allocation rules corresponding to paths on a finite grid, i.e., average of A^P for all $P \in \mathcal{P}^n$ for some n. Then, we let the grid size converge to zero. We will show that the limit of the corresponding allocation rules exists and that it is a well-defined value also in case of non-differentiability of the risk capital function r. We show that in case of differentiability the outcome is the standard Aumann-Shapley value. In case of non-differentiability however, the outcome can be regarded as a weighted average of standard Aumann-Shapley values for "nearby" risk capital allocation problems with a differentiable risk capital function. Moreover we will provide a geometric interpretation of the corresponding weights. We will refer to this value as the *Weighted-Aumann Shapley value*.

In Definition 5.2 we introduced an allocation rule based on a path $P \in \mathcal{P}^n$. Since A^P is a risk capital allocation rule (see Proposition 5.3), the average of all path based risk capital allocation rules is a risk capital allocation rule itself. This allocation rule is defined as follows.

Definition 6.1 *Let $n \in \mathbb{N}$. Then, $K^n : \mathcal{R} \to \mathbb{R}^N$ is defined by*

$$K^n(R) = \frac{1}{|\mathcal{P}^n|} \cdot \sum_{P \in \mathcal{P}^n} A^P(R), \qquad \textit{for all } R \in \mathcal{R}, \tag{40}$$

where A^P is defined in Definition 5.2.

For a given n, one can rewrite the definition of K^n to the standard Aumann-Shapley method, which is proposed by Moulin (1995) for a given discrete production problem with a continuously differentiable production function. If $n = 1$, this allocation rule equals the Shapley value, i.e., $K^1 = S$. The asymptotic behavior of $K^n(R)$ when $n \to \infty$ (or, equivalently, when the grid size converges to 0) is a central topic of this paper. Next, we rewrite $K^n(R)$ as a weighted sum of marginal contributions over all participation profiles on the grid G^n.

Proposition 6.2 *Let $R \in \mathcal{R}$ and $n \in \mathbb{N}$. Then, we have for all $i \in N$ that*

$$K_i^n(R) = \sum_{\lambda \in G^n : \lambda_i < 1} t^n(\lambda) \cdot p_i^n(\lambda) \cdot [r(\lambda + (1/n) \cdot e_i) - r(\lambda)], \tag{41}$$

where

$$t^n(\lambda) = \frac{\prod_{j \in N} \begin{pmatrix} n \\ n\lambda_j \end{pmatrix}}{\begin{pmatrix} |N|n \\ |N|n\bar{\lambda} \end{pmatrix}}, \tag{42}$$

and

$$p_i^n(\lambda) = \frac{1 - \lambda_i}{\sum_{j \in N}(1 - \lambda_j)}, \tag{43}$$

for all $\lambda \in G^n \backslash \{e_N\}$,

$$\bar{\lambda} = \frac{1}{|N|} \sum_{i \in N} \lambda_i, \qquad \textit{for all } \lambda \in \mathbb{R}^N, \tag{44}$$

and where the risk capital function r is defined in (17).

Proof We provide the proof in Appendix A. □

The function $t^n(\lambda)$ represents the probability that λ lies on a path, if we randomly select a path from $\mathcal{P}^n$ according to the discrete uniform distribution. Moreover, $p_i^n(\lambda)$ is the conditional probability that $\lambda + (1/n) \cdot e_i$ lies on a path, provided that the path passes through λ. So, in order to compute $K_i^n(R)$, each marginal contribution $r(\lambda + (1/n) \cdot e_i) - r(\lambda)$ is multiplied by the probability that both λ and $\lambda + (1/n) \cdot e_i$ are on a path.

We will show in the sequel that $\lim_{n\to\infty} K^n(R)$ exists if $R \in \mathcal{R}$. This enables us to define the Weighted-Aumann Shapley value $WAS : \mathcal{R} \to \mathbb{R}^N$ by

$$WAS(R) = \lim_{n \to \infty} K^n(R), \qquad \text{for all } R \in \mathcal{R}. \tag{45}$$

Moreover, we will show that this allocation rule satisfies $WAS(R) = AS(R)$ in case the risk capital function r is partially differentiable in e_N. So, K^1 is the Shapley value and $\lim_{n\to\infty} K^n$ the Aumann-Shapley value, if existent.

Moreover, in case of non-differentiability, we show that the Weighted Aumann-Shapley value is a weighted average of standard Aumann-Shapley values of "nearby" risk capital allocation problems where the risk capital function is partially differentiable.

6.2 Convergence

In this subsection, we provide our main result of this paper. We show that the Weighted Aumann-Shapley value exists and provide a closed form solution. We use the following notation throughout this subsection:

- We use the Bachmann-Landau notation. Let $f, g : \mathbb{N} \to \mathbb{R}$ be two real-valued functions. Then, we write $f(n) = \mathcal{O}(g(n))$ if there is a $K > 0$ such that $|f(n)| \leq K \cdot |g(n)|$ for every $n \in \mathbb{N}$. If $f : \mathbb{N} \to \mathbb{R}$ is such that $f(n) = \mathcal{O}(n^{-p})$ for every $p > 0$, we write $f(n) = \mathcal{O}(n^{-\infty})$. Moreover, if $g : \mathbb{R}_{++} \to \mathbb{R}$ is such that there is a $K > 0$ such that $|g(\varepsilon)| \leq K \cdot \varepsilon$ for every $\varepsilon > 0$, we write $g(\varepsilon) = \mathcal{O}(\varepsilon)$. Here, $\mathbb{R}_{++} = (0, \infty)$ is the set of all positive, real numbers.

- Let $f : \mathbb{R}_{++} \times \mathbb{N} \to \mathbb{R}$ and $g : \mathbb{N} \to \mathbb{R}$. Then, we write $f(\varepsilon, n) = \mathcal{O}^{\varepsilon}(g(n))$ if for every $\varepsilon > 0$, there is a $K_\varepsilon > 0$ such that $|f(\varepsilon, n)| \leq K_\varepsilon \cdot |g(n)|$ for all $n \in \mathbb{N}$. This notation is an extension of the standard Bachmann-Landau notation.

- For all $\lambda \in \mathbb{R}^N$, we write $\|\lambda\| = \sqrt{\sum_{i\in N} \lambda_i^2}$ as the Euclidean norm of λ.
- We define the set of participation profiles that are not nearby $\lambda = e_\emptyset$ and e_N as follows. For all $n \in \mathbb{N}$ and $\varepsilon > 0$, we define

$$G_\varepsilon = \left\{\lambda \in [0,1]^N : \varepsilon \leq \bar{\lambda} \leq 1-\varepsilon\right\},$$

  and

$$G_\varepsilon^n = G^n \cap G_\varepsilon,$$

  where $\bar{\lambda}$ is defined in (44).
- We define D^d as the set of participation profiles in the d-environment of the diagonal, i.e., for all $d > 0$, we have

$$D^d = \{\lambda \in [0,1]^N : \|\lambda - \bar{\lambda}\cdot e_N\| < d\}.$$

  Moreover, we define for all $n \in \mathbb{N}$ the set

$$D(n) = D^{d_n}, \qquad \text{where } d_n = n^{-\frac{1}{2}+\frac{1}{8|N|}}.$$

In Figure 3, we provide an illustration of th sets G_ε and D^d in case of two divisions. We will only consider participation profiles in G_ε for an arbitrary choice of $\varepsilon > 0$. In the following proposition, we approximate the weight functions t^n and p_i^n that are defined in Proposition 6.2. For large n, we obtain expressions with a nice interpretation, which are obtained using, among others, several Taylor and Stirling approximations. We get that $t^n(\lambda)$ is exponentially small for λ away from the diagonal.

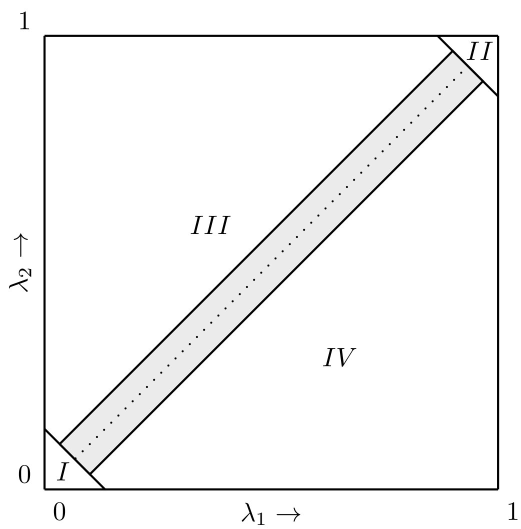


Figure 3: The shaded set is a set of participation profiles with substantial aggregate contribution to the Weighted Aumann-Shapley value (see Lemma C.1 and Lemma C.2) in case $|N| = 2$. Here, $I \cup II = [0,1]^N \backslash G_\varepsilon$ and $III \cup IV = G_\varepsilon \backslash D(n)$ for an arbitrary choice of $\varepsilon > 0$ and $n \in \mathbb{N}$.

Proposition 6.3 *Let $i \in N$ and define $Dom = \{(\varepsilon, n, \lambda) : \varepsilon > 0, n \in \mathbb{N}, \lambda \in G_\varepsilon^n\}$. Then, we have*

$$t^n(\lambda) = \begin{cases} \left(e^{-c(\bar{\lambda})\cdot n\cdot\|\lambda-\bar{\lambda}\cdot e_N\|^2}\right)\cdot b(n,\bar{\lambda})\cdot[1+\mathcal{O}^\varepsilon(n^{-\frac{1}{4}})], & \text{if } \lambda \in D(n), \quad (46)\\ \mathcal{O}^\varepsilon(n^{-\infty}), & \text{if } \lambda \notin D(n), \quad (47)\end{cases}$$

and

$$p_i^n(\lambda) = \begin{cases} \dfrac{1}{|N|} \cdot [1 + \mathcal{O}^\varepsilon(n^{-\frac{1}{4}})], & \text{if } \lambda \in D(n), \quad (48) \\ \mathcal{O}(1), & \text{if } \lambda \notin D(n), \quad (49) \end{cases}$$

for all $(\varepsilon, n, \lambda) \in Dom$, *where*

$$c(\bar\lambda) = \frac{1}{2\bar\lambda(1-\bar\lambda)} > 0, \tag{50}$$

and

$$b(n, \bar\lambda) = (2\pi n)^{\frac{1}{2}(1-|N|)} \cdot \sqrt{|N|} \cdot (\bar\lambda(1-\bar\lambda))^{\frac{1}{2}(1-|N|)}. \tag{51}$$

Proof We provide the proof in Appendix B. □

For large n, we get that $t^n(\lambda)$ only depends on λ via $\bar\lambda$ and $\|\lambda - \bar\lambda \cdot e_N\|$ and that $p_i^n(\lambda)$ is symmetric close to the diagonal. For a given $n \in \mathbb{N}$ and $\bar\lambda \in \{0, \frac{1}{n}, \ldots, 1\}$, the function $b(n, \bar\lambda)$ is approximately the probability that a path goes through the diagonal (i.e., through $\bar\lambda \cdot e_N$) and $c(\bar\lambda)$ indicates a speed at which $t^n(\lambda)$ converges to zero for participation profiles away from the diagonal. The function $t^n(\lambda)$ is exponentially small in n if λ is not nearby to the diagonal, i.e., $\lambda \notin D(n)$. Moreover, $p_i^n(\lambda)$ is bounded. Therefore, only participation profiles very close to the diagonal are relevant for K^n if n converges to infinity. From Proposition 6.3, we obtain that $t^n(\lambda) \cdot p_i^n(\lambda)$ can be approximated for λ nearby the diagonal using the following formula.

Definition 6.4 *The function* $h^n : [0,1]^N \backslash \{e_\emptyset, e_N\} \to \mathbb{R}_{++}$ *is given by*

$$h^n(\lambda) = \left(e^{-c(\bar\lambda)\cdot n \cdot \|\lambda - \bar\lambda \cdot e_N\|^2}\right) \cdot b(n, \bar\lambda) \cdot \frac{1}{|N|},$$

for all $\lambda \in [0,1]^N \backslash \{e_\emptyset, e_N\}$ *and* $n \in \mathbb{N}$, *where* $c(\bar\lambda)$ *is defined in (50) and* $b(n, \bar\lambda)$ *in (51).*

From Proposition 6.3, we get that

$$t^n(\lambda) \cdot p_i^n(\lambda) = h^n(\lambda) \cdot [1 + \mathcal{O}^\varepsilon(n^{-\frac{1}{4}})], \tag{52}$$

for all $(\varepsilon, n, \lambda) \in Dom$ such that $\lambda \in D(n)$.

In Lemma C.1 and Lemma C.2, we show that all participation profiles in G^n that are not close to the diagonal or that are nearby $e_\emptyset$ or e_N have a negligible aggregate contribution to K^n if n converges to infinity. In case of two divisions, we illustrate these participation profiles in Figure 3. We obtain that all participation profiles in K_m for $m \notin \{1, \ldots, p^*\}$ have a negligible aggregate contribution.

From Corollary 4.7, we get that the risk capital function r is almost everywhere partially differentiable. In Lemma C.6, we extent this result by showing that participation profiles in a $\frac{1}{n}$-environment of participation profiles where r is non-differentiable have a negligible aggregate contribution for large n as well. For all other risk profiles λ, we obtain from (24) that

$$r(\lambda + (1/n) \cdot e_i) - r(\lambda) = \frac{1}{n} \cdot E_{\mathbb{Q}_m}[X_i],$$

for all $i \in N$ and $\lambda \in G^n \cap K_m$ such that $\lambda_i < 1$ and $m \in \{1, \ldots, p\}$. So, for these risk profiles, we know the marginal effect exactly.

Proposition 6.5 *Let* $R \in \mathcal{R}$. *Then, for all* $i \in N$ *we have*

$$K_i^n(R) = \sum_{m=1}^{p^*} E_{\mathbb{Q}_m}[X_i] \cdot \phi_m^{n,\varepsilon} + \mathcal{O}(\varepsilon) + \mathcal{O}^\varepsilon(n^{-\frac{1}{4}}),$$

where

$$\phi_m^{n,\varepsilon} = \frac{1}{n} \cdot \sum_{\lambda \in G_\varepsilon^n \cap D(n) \cap K_m} h^n(\lambda), \tag{53}$$

and p^*, $\mathbb{Q}_m$ and K_m are defined in Definition 4.3.

Proof We provide the proof in Appendix C. □

The expression $\phi_m^{n,\varepsilon}$ is a weight for a gradient of the risk capital function r "nearby" the diagonal, namely $(E_{\mathbb{Q}_m}[X_i])_{i\in N}$. Next, we show that we can replace this weight by an expression that has a geometric interpretation and is not dependent on n or ε anymore. This result is obtained by replacing the sum in (53) by an integral (see Lemma D.2 and Lemma D.3) and, thereafter, solving this integral.

Proposition 6.6 *For all $R \in \mathcal{R}$, it holds that*

$$K_i^n(R) = \sum_{m=1}^{p^*} E_{\mathbb{Q}_m}[X_i] \cdot \phi_m + \mathcal{O}(\varepsilon) + \mathcal{O}^\varepsilon(n^{-\frac{1}{4}}), \qquad \text{for all } i \in N,$$

where for all $m \in \{1, \ldots, p^*\}$ we define

$$S = \left\{ z \in \mathbb{R}^N : \sum_{i \in N} z_i = 0, \|z\| = 1 \right\}$$

$$S_m = \left\{ z \in S : f_{\mathbb{Q}_m}(z) = \max_{\ell \in \{1,\ldots,p^*\}} f_{\mathbb{Q}_\ell}(z) \right\},$$

and

$$\phi_m = \frac{\mu(S_m)}{\mu(S)},$$

where μ is the surface area measure on S and the function $f_{\mathbb{Q}}$ is defined in Definition 4.1.

Proof We provide the proof in Appendix D. □

Remark that from Lemma 4.6, we get

$$\sum_{m=1}^{p^*} \phi_m = 1. \tag{54}$$

Recall the definition of the Weighted Aumann-Shapley value in (45). Using Proposition 6.6, we next show that this value exists and, moreover, we provide a closed form expression. This result is given in the following theorem.

Theorem 6.7 *For all $R \in \mathcal{R}$, we have that $WAS(R) = \lim_{n\to\infty} K^n(R)$ exists and, moreover, it is given by*

$$WAS_i(R) = \sum_{m=1}^{p^*} E_{\mathbb{Q}_m}[X_i] \cdot \phi_m, \qquad \text{for all } i \in N,$$

with ϕ_m as defined in Proposition 6.6.

Proof Let $R \in \mathcal{R}$. From Proposition 6.6, we get for all $n \in \mathbb{N}$ and $\varepsilon > 0$ that

$$\left| K_i^n(R) - \sum_{m=1}^{p^*} E_{\mathbb{Q}_m}[X_i] \cdot \phi_m \right| < K \cdot \varepsilon + L_\varepsilon \cdot n^{-\frac{1}{4}}, \qquad \text{where } K, L_\varepsilon > 0.$$

Pick an $\eta > 0$. Let $\varepsilon = \frac{\eta}{2K}$ and N_η such that $L_\varepsilon \cdot N_\eta^{-\frac{1}{4}} = \frac{1}{2}\eta$. Then, we have for all $n > N_\eta$ that

$$\left| K_i^n(R) - \sum_{m=1}^{p^*} E_{\mathbb{Q}_m}[X_i] \cdot \phi_m \right| < \eta.$$

This concludes the proof. □

Theorem 6.7 shows that the Weighted Aumann-Shapley value is a convex combination of regular Aumann-Shapley values of "nearby" risk capital allocation problems. The weight ϕ_m has a geometric interpretation, as we will show in the next constructive example.

Example 6.8 *In this example, we discuss a case with three divisions. Let $N = \{1, 2, 3\}$ $\Omega = \{\omega_1, \dots, \omega_5\}$ and $\mathbb{P}(\{\omega\}) = \frac{1}{5}$ for all $\omega \in \Omega$. Moreover, let the risk capital allocation problem given by $R = ((X_i)_{i \in N}, \rho_{0.1}^{ES}) \in \mathcal{R}$ such that*

$$X_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \\ -1 \end{bmatrix}, \; X_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \\ 2 \end{bmatrix}, \; \textit{and } X_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \\ -1 \end{bmatrix}.$$

Next, we show the value of ϕ_m. We pick out the colored triangle $T := \{\lambda \in [0,1]^N : \bar{\lambda} = \frac{1}{3}\}$ as in Figure 4.

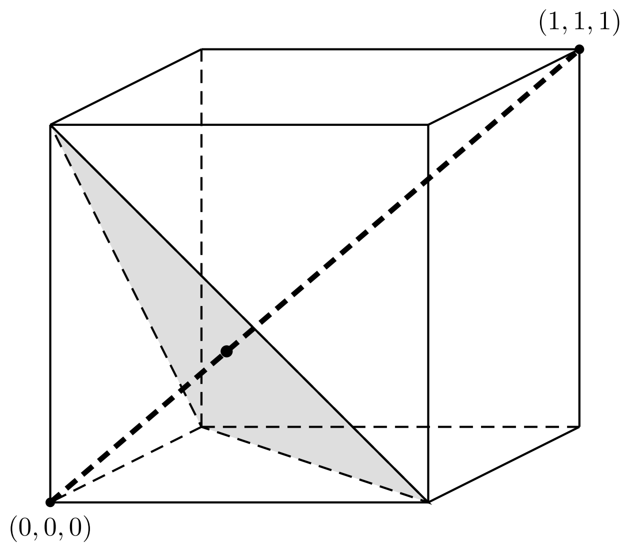
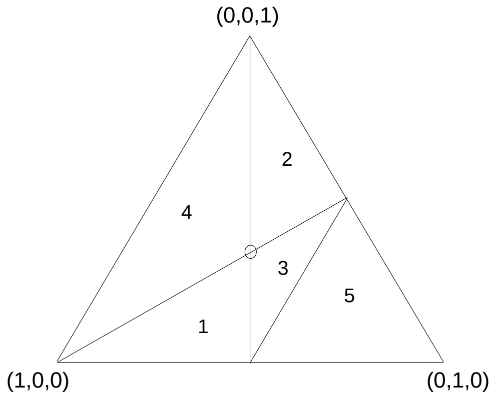


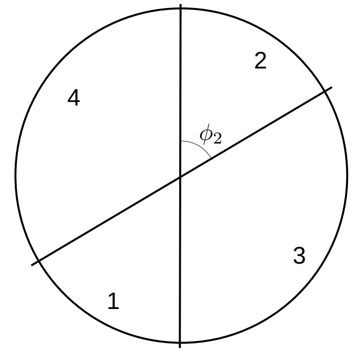
Figure 4: A representation of the path via the diagonal for $N = \{1, 2, 3\}$ in $[0,1]^N$. The colored area is the set $T := \{\lambda \in [0,1]^N : \bar{\lambda} = \frac{1}{3}\}$.

The set $T \backslash L(R)$ is a union of line segments on the triangle T. This set contains all participation profiles in T where the risk capital function r is non-differentiable. Let $\mathbb{Q}_m = (1_{\omega_m}, 0_{-\omega_m})$ for $m \in \{1, \dots, 5\}$. We display the set $(T \cap K_m)_{m \in \{1,\dots,p\}}$ in Figure 5.a and we immediately see that $p^* = 4$. Then, the fraction ϕ_m corresponds with the normalized angle that the set $T \cap K_m$ form in point $\lambda = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$. We depict a sufficiently small environment $T \cap D^d$ around this point (as shown in Figure 5.a) and display this set in Figure 5.b. We obtain $\phi_1 = \frac{1}{6}$, $\phi_2 = \frac{1}{6}$, $\phi_3 = \frac{1}{3}$ and $\phi_4 = \frac{1}{3}$. Note that, instead of the triangle T with $\bar{\lambda} = \frac{1}{3}$, we could have depicted every triangle where $0 < \bar{\lambda} < 1$. Then, we obtain from Theorem 6.7 that

$$WAS(R) = \left(\frac{1}{2}, 0, \frac{1}{2}\right).$$



(a) Simplex T of three-division unit-cube. The numbers reflect the areas: $m : T \cap K_m$ for all $m \in \{1, 2, 3, 4, 5\}$.



(b) A sufficiently small environment of the diagonal, where ϕ_2 corresponds with the normalized angle of $T \cap K_2$ in $\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$.

Figure 5: Illustration of the Weighted Aumann-Shapley value corresponding to Example 6.8.

So, Division 2, which holds the portfolio with the highest expected loss, get assigned the lowest risk capital allocation. This is due to large hedge benefits as the loss in state ω_5 is high for this portfolio, while this state is the best case scenario for the firm. $\bigtriangledown$

Remark We can prove Theorem 6.7 using a diagonal width $d_n = n^{-\frac{1}{2}+\delta}$ for all $\delta \in \left(0, \frac{1}{2(|N|+2)}\right)$. Then, we can adjust Proposition 6.6 as follows:

$$K_i^n(R) = \sum_{m=1}^{p^*} E_{\mathbb{Q}_m}[X_i] \cdot \phi_m + \mathcal{O}(\varepsilon) + \mathcal{O}^{\varepsilon}(n^{-\frac{1}{2}+\delta(|N|+2)}), \qquad \text{for all } R \in \mathcal{R} \text{ and } i \in N.$$

In this section, we depicted $\delta = \frac{1}{8|N|}$ to avoid tedious results.

6.3 Properties of the Weighted Aumann-Shapley value

In this subsection, we show some properties of the Weighted Aumann-Shapley value. As stated earlier, the Weighted Aumann-Shapley value equals the Aumann-Shapley if the risk capital function r is differentiable in e_N. This follows directly from Theorem 6.7 and is shown in the following corollary.

Corollary 6.9 *For all $R \in \mathcal{R}'$, it holds that*

$$WAS(R) = AS(R).$$

Proof For all $R \in \mathcal{R}'$, we have that the risk capital function r is partially differentiable in e_N and, so, $p^* = 1$. Then, we get

$$\begin{aligned} WAS(R) &= (E_{\mathbb{Q}_1}[X_i])_{i \in N} && (55)\\ &= AS(R), && (56) \end{aligned}$$

where (55) follows from Theorem 6.7 and (54), and (56) follows from (20). This concludes the proof. $\square$

This result is stated for the Aumann-Shapley mechanism in cost sharing by Moulin (1995). Moulin, however, typically requires the function r to be continuously differentiable, whereas this typically does not need to be satisfied in case that ρ is a coherent risk measure.

We next show a range of values of ϕ_m.

Proposition 6.10 *For all $R \in \mathcal{R}$ and $m \in \{1, \ldots, p^*\}$, we have*

$$\phi_m \in \left[0, \frac{1}{2}\right] \cup \{1\}.$$

Proof Let $R \in \mathcal{R}$. Suppose $\phi_m > \frac{1}{2}$ for an $m \in \{1, \ldots, p^*\}$ and $p^* > 1$. Note that $\mu(S_\ell \cap S_{\ell'}) = 0$ for all $\ell, \ell' \in \{1, \ldots, p^*\}$ such that $\ell \neq \ell'$, where μ is the surface area on S. Moreover, we have $z \in S$ if and only if $-z \in S$. From these two results, we obtain that there must exists a $z \in S_m$ such that $f_{\mathbb{Q}_m}(z) > \max_{\ell \in \{1,\ldots,p^*\}\setminus m} f_{\mathbb{Q}_\ell}(z)$ and $-z \in S_m$. So, let $z \in S_m$ such that $f_{\mathbb{Q}_m}(z) > \max_{\ell \in \{1,\ldots,p^*\}\setminus m} f_{\mathbb{Q}_\ell}(z)$. By linearity of $f_{\mathbb{Q}}$, we get $f_{\mathbb{Q}}(-z) = -f_{\mathbb{Q}}(z)$. If $-z \in S_m$, we get $f_{\mathbb{Q}_m}(-z) \geq \max_{\ell \in \{1,\ldots,p^*\}\setminus m} f_{\mathbb{Q}_\ell}(-z)$ or, equivalently,

$$f_{\mathbb{Q}_m}(z) \leq \min_{\ell \in \{1,\ldots,p^*\}\setminus m} f_{\mathbb{Q}_\ell}(z),$$

which is a contradiction. Hence, we have $\phi_m \leq \frac{1}{2}$ or $p^* = 1$ so that $\phi_m = 1$. This concludes the result. □

The range of ϕ_m in Proposition 6.10 is tight for $|N| > 2$, i.e., for every $c \in [0, \frac{1}{2}] \cup \{1\}$, one can construct a risk capital allocation problem such that $\phi_m = c$ for an $m \in \{1, \ldots, p^*\}$.

Next, we generalize the property that the Aumann-Shapley value, if existent, is in the single-valued fuzzy core (see Theorem 3.4). We namely show that the Weighted Aumann-Shapley is always is the fuzzy core.

Proposition 6.11 *The Weighted Aumann-Shapley value satisfies Fuzzy Core Selection on $\mathcal{R}$.*

Proof Let $R \in \mathcal{R}$. It holds that

$$\begin{aligned} FCore(R) &= \text{conv}\{(E_{\mathbb{Q}}[X_i])_{i \in N} : \mathbb{Q} \in Q^*(\rho)\} && (57) \\ &= \text{conv}\{(E_{\mathbb{Q}_m}[X_i])_{i \in N} : m \in \{1, \ldots, p^*\}\}, && (58) \end{aligned}$$

where $Q^*(\rho)$ is defined in (19). Here, (57) follows from Theorem 3.5 and (58) follows from Definition 4.3. Moreover, we showed in Theorem 6.7 that $WAS(R)$ is given by a convex combination of allocation $(E_{\mathbb{Q}_m}[X_i])_{i \in N}$ for all $m \in \{1, \ldots, p^*\}$. Hence, we have $WAS(R) \in FCore(R)$. □

All properties that are satisfied by path based allocation rules are also satisfied by the Weighted Aumann-Shapley value. This holds as this rule is the average over allocation rules based on a path. Hence, according to Theorem 5.4, the Weighted Aumann-Shapley value satisfies *Translation Invariance*, *Scale Invariance* and *Monotonicity* on $\mathcal{R}$.

Let $\nabla^+ r(e_N)$ be the right derivative of the risk capital function r in e_N and $\nabla^- r(e_N)$ the corresponding left derivative. We straightforwardly obtain from Theorem 6.7 the following result for risk capital allocation problems with two divisions.

Corollary 6.12 *If $|N| = 2$ and $R \in \mathcal{R}$, we have*

$$WAS(R) = \frac{1}{2}\nabla^+ r(e_N) + \frac{1}{2}\nabla^- r(e_N).$$

One can easily verify that $\nabla^+ r(e_N) = (\max\{E_{\mathbb{Q}}[X_i] : \mathbb{Q} \in Q^*(\rho)\})_{i \in N}$ and $\nabla^- r(e_N) = (\min\{E_{\mathbb{Q}}[X_i] : \mathbb{Q} \in Q^*(\rho)\})_{i \in N}$, where $Q^*(\rho)$ is defined in (19).

Remark We note that we only need positive homogeneity and piecewise linearity of the risk capital function r for Theorem 6.7 to hold, where $E_{\mathbb{Q}_m}[X_i]$ is then replaced by $\frac{\partial r}{\partial \lambda_i}(\lambda)$ for $\lambda \in K_m$ with obvious meaning of notation. On this class of allocation problems, also Proposition 6.10, Proposition 6.11 and Corollary 6.12 hold. For only a piecewise linear function r, one can also deduce a closed form expression of K^n in Theorem 6.7 based on derivatives of r in participation profiles along the diagonal.

7 Conclusion

This paper considers the allocation problem that arises when the total risk capital withheld by a firm needs to be divided over several portfolios within the firm. We propose a generalization of the Aumann-Shapley value for risk capital allocation problems. The Aumann-Shapley value receives considerable attention in the literature. It is well-known that this allocation rule requires differentiability of the fuzzy game for existence.

Our rule is also well-defined in case a differentiability condition is not satisfied. We introduce this allocation rule inspired by the Shapley value in a fuzzy setting. It follows a much weaker asymptotic approach than the one proposed by Aumann and Shapley (1974). The asymptotic approach of Aumann and Shapley (1974) is not valid for fuzzy games related to risk capital allocation problems. We take a grid on a fuzzy participation set, define paths on this grid and construct an allocation rule based on a path. Then, we show that the limit of the average over these allocations exists, when the grid size converges to zero. We define the Weighted Aumann-Shapley value as this limit. We provide an explicit formula for this allocation rule, which has a geometric interpretation. This allocation rule happens to coincide with the Mertens value. Moreover, it satisfies some properties which are known to hold for the Aumann-Shapley value.

References

Acerbi, C. and D. Tasche (2002). On the coherence of expected shortfall. *Journal of Banking and Finance 26*, 1487–1503.

Artzner, P., F. Delbaen, J. Eber, and D. Heath (1999). Coherent measures of risk. *Mathematical Finance 9*, 203–228.

Aubin, J. (1979). *Mathematical Methods of Game and Economic Theory*. Amsterdam: North-Holland.

Aubin, J. (1981). Cooperative fuzzy games. *Mathematics of Operations Research 6*, 1–13.

Aumann, R. and L. Shapley (1974). *Values of Non-Atomatic Games*. Princeton: Princeton University Press.

Billera, L. J. and D. Heath (1982). Allocation of shared costs: A set of axioms yielding a unique procedure. *Mathematics of Operations Research 7*, 32–39.

Delbaen, F. (2000). Coherent risk measures on general probability spaces. In *Avances in Finance and Stochastics, Essays in Honour of Dieter Sondermann*. New York: Springer.

Denault, M. (2001). Coherent allocation of risk capital. *Journal of Risk 4*, 1–34.

Denneberg, D. (1994). *Non-additive measure and integral*. Dordrecht: Kluwer Academic Publishers.

Dhaene, J., M. Goovaerts, and R. Kaas (2003). Economic capital allocation derived from risk measures. *North American Actuarial Journal 7*, 44–59.

Dhaene, J., S. Vanduffel, M. Goovaerts, R. Kaas, Q. Tang, and D. Vyncke (2006). Risk measures and comonotonicity: A review. *Stochastic Models 22*, 573–606.

Gulick, G. van, A. De Waegenaere, and H. Norde (2012). Excess based allocation of risk capital. *Insurance: Mathematics and Economics 50*, 26–42.

Haimanko, O. (2001). Cost sharing: the nondifferentiable case. *Journal of Mathematical Economics 35*, 445–462.

Kalkbrener, M. (2005). An axiomatic approach to capital allocation. *Mathematical Finance 15*, 425–437.

Laeven, R. and M. Goovaerts (2004). An optimization approach to the dynamic allocation of economic capital. *Insurance: Mathematics and Economics 35*, 299–319.

Mertens, J. (1980). Values and derivatives. *Mathematics of Operations Research 5*, 523–552.

Mertens, J. (1988). The Shapley value in the non differentiable case. *International Journal of Game Theory 17*, 1–65.

Mirman, L. and I. Tauman (1982). Demand compatible equitable cost sharing prices. *Mathematics of Operations Research 7*, 40–56.

Moulin, H. (1995). On additive methods to share joint costs. *Japanese Economic Review 46*, 303–332.

Myers, S. C. and J. A. Read (2001). Capital allocation for insurance companies. *Journal of Risk and Insurance 68*, 545–580.

Neyman, A. (2002). Values of games with infinitely many players. Volume 3 of *Handbook of Game Theory with Economic Applications*, pp. 2121 – 2167. Elsevier.

Neyman, A. and R. Smordinsky (2004). Asymptotic values of vector measure games. *Mathematics of Operations Research 29*, 739–775.

Schmeidler, D. (1986). Integral representation without additivity. *Proceedings of the American Mathematical Society 97*, 255–261.

Shapley, L. (1953). A value for n-person games. In *Contributions to the Theory of Games*, pp. 307–317. Princeton University Press.

Shapley, L. (1971). Cores of convex games. *International Journal of Game Theory 1*, 11–26.

Sprumont, Y. (2005). On the discrete version of the Aumann-Shapley cost-sharing method. *Econometrica 73*, 1693–1712.

Tasche, D. (1999). Risk contributions and performance measurement. Preprint, TU Munich.

Tasche, D. (2002). Expected shortfall and beyond. *Journal of Banking and Finance 26*, 1519–1533.

Tsanakas, A. (2004). Dynamic capital allocation with distortion risk measures. *Insurance: Mathematics and Economics 35*, 223–243.

Tsanakas, A. (2009). To split or not to split: Capital allocation with convex risk measures. *Insurance: Mathematics and Economics 44*, 268–277.

Tsanakas, A. and C. Barnett (2003). Risk capital allocation and cooperative pricing of insurance liabilities. *Insurance: Mathematics and Economics 33*, 239–254.

Wang, S. S. (1995). Insurance pricing and increased limits ratemaking by proportional hazards transforms. *Insurance: Mathematics and Economics 17*, 43–54.

Wang, S. S., V. R. Young, and H. H. Panjer (1997). Axiomatic characterization of insurance prices. *Insurance: Mathematics and Economics 21*, 173–183.

Wang, Y. (1999). The additivity and dummy axioms in the discrete cost sharing model. *Economic Letters 64*, 187–192.

A Proof of Proposition 6.2

In this appendix, we use the following notation. The set $\widetilde{G}_k^n$ is given by

$$\widetilde{G}_k^n = \left\{ \lambda \in G^n : \sum_{i \in N} \lambda_i = \frac{k}{n} \right\}, \qquad \text{for all } n \in \mathbb{N} \text{ and } k \in \{0, \ldots, |N|n\}. \tag{59}$$

The set $\widetilde{G}_k^n$ consists of all participation profiles on the grid where the sum of the coordinates is constant. Note that we have

$$\widetilde{G}_k^n = \{P(k) : P \in \mathcal{P}^n\}, \qquad \text{for all } n \in \mathbb{N} \text{ and } k \in \{0, \ldots, |N|n\}. \tag{60}$$

Next, we show (41). Expression (40) of $K^n(R)$ can be rewritten as

$$\begin{aligned}
K^n(R) &= \frac{1}{|\mathcal{P}^n|} \cdot \sum_{P \in \mathcal{P}^n} A^P(R) &(61)\\
&= \frac{1}{|\mathcal{P}^n|} \cdot \sum_{P \in \mathcal{P}^n} \sum_{k=0}^{|N|n-1} [r(P(k+1)) - r(P(k))] \cdot e_{i(P,k)} &(62)\\
&= \sum_{k=0}^{|N|n-1} \sum_{P \in \mathcal{P}^n} \frac{1}{|\mathcal{P}^n|} \cdot [r(P(k+1)) - r(P(k))] \cdot e_{i(P,k)}, &(63)
\end{aligned}$$

where (61) follows from Definition 6.1 and (62) follows from Definition 5.2. Let $i \in N$. Then, we obtain

$$\begin{aligned}
K_i^n(R) &= \sum_{k=0}^{|N|n-1} \sum_{P \in \mathcal{P}^n : i(P,k)=i} \frac{1}{|\mathcal{P}^n|} \cdot [r(P(k+1)) - r(P(k))] &(64)\\
&= \sum_{k=0}^{|N|n-1} \sum_{P \in \mathcal{P}^n : i(P,k)=i} \frac{1}{|\mathcal{P}^n|} \cdot [r(P(k) + (1/n) \cdot e_i) - r(P(k))] &(65)\\
&= \sum_{k=0}^{|N|n-1} \sum_{\lambda \in \widetilde{G}_k^n} \sum_{\substack{P \in \mathcal{P}^n : \\ i(P,k)=i, P(k)=\lambda}} \frac{1}{|\mathcal{P}^n|} \cdot [r(\lambda + (1/n) \cdot e_i) - r(\lambda)] &(66)\\
&= \sum_{k=0}^{|N|n-1} \sum_{\lambda \in \widetilde{G}_k^n : \lambda_i < 1} [r(\lambda + (1/n) \cdot e_i) - r(\lambda)] \cdot \sum_{\substack{P \in \mathcal{P}^n : \\ i(P,k)=i, P(k)=\lambda}} \frac{1}{|\mathcal{P}^n|} &(67)\\
&= \sum_{k=0}^{|N|n-1} \sum_{\lambda \in \widetilde{G}_k^n : \lambda_i < 1} [r(\lambda + (1/n) \cdot e_i) - r(\lambda)] \cdot t^n(\lambda) \cdot p_i^n(\lambda) &(68)\\
&= \sum_{\lambda \in G^n : \lambda_i < 1} [r(\lambda + (1/n) \cdot e_i) - r(\lambda)] \cdot t^n(\lambda) \cdot p_i^n(\lambda), &(69)
\end{aligned}$$

where we define

$$t^n(\lambda) = \frac{\left|\left\{P \in \mathcal{P}^n : P\left(|N|n\bar{\lambda}\right) = \lambda\right\}\right|}{|\mathcal{P}^n|},$$

as the fraction of paths in $\mathcal{P}^n$ that pass through λ and

$$p_i^n(\lambda) = \frac{\left|\left\{P \in \mathcal{P}^n : P\left(|N|n\bar{\lambda}\right) = \lambda, i\left(P, |N|n\bar{\lambda}\right) = i\right\}\right|}{\left|\left\{P \in \mathcal{P}^n : P\left(|N|n\bar{\lambda}\right) = \lambda\right\}\right|},$$

as the fraction of the paths in $\mathcal{P}^n$ passing through λ, that pass through $\lambda + \frac{1}{n} \cdot e_i$ as well. Here, (64) follows from (63), (65) follows from (28), (66) follows from (60), (67) follows from the fact that if $k \in \{0, \ldots, |N|n-1\}$ and $\lambda \in \widetilde{G}_k^n$ are such that $\lambda_i = 1$ then no path $P \in \mathcal{P}^n$ exists with $i(P,k) = i$ and $P(k) = \lambda$, (68) follows from the fact that if $k \in \{0, \ldots, |N|n-1\}$ and $\lambda \in \widetilde{G}_k^n$ are such that $P(k) = \lambda$ then $k = |N|n\bar{\lambda}$ and (69) follows from the fact that $\bigcup_{k=1}^{|N|n-1} G_k^n = G^n$ and $G_{k_1}^n \cap G_{k_2}^n = \emptyset$ if $k_1 \neq k_2$.

Next, we show (42). Any path can be regarded as an ordered sequence of $|N|n$ steps, where for every division $i \in N$ precisely n steps are made in the direction of division i. Hence,

$$|\mathcal{P}^n| = \frac{(|N|n)!}{(n!)^{|N|}}. \tag{70}$$

Let $\lambda \in G^n \backslash \{e_N\}$. The number of paths P in $\mathcal{P}^n$ such that $P\left(|N|n\bar{\lambda}\right) = \lambda$ is given by

$$\left|\left\{P \in \mathcal{P}^n : P\left(|N|n\bar{\lambda}\right) = \lambda\right\}\right| = \frac{\left(|N|n\bar{\lambda}\right)! \cdot \left(|N|n(1-\bar{\lambda})\right)!}{\prod_{j \in N}(n\lambda_j)! \cdot (n(1-\lambda_j))!}. \tag{71}$$

Hence, one can verify that dividing (71) by (70) yields (42). Note that, keeping $\bar{\lambda}$ constant, the various values of $t^n(\lambda)$ constitute a density function of some multivariate hypergeometric distribution.

Finally, we show (43). The number of paths P in $\mathcal{P}^n$ with $P(|N|n\bar{\lambda}) = \lambda$ and $i(P, |N|n\bar{\lambda}) = i$ (i.e. passing through λ and $\lambda + (1/n) \cdot e_i$) is given by:

$$\left|\left\{P \in \mathcal{P}^n : P\left(|N|n\bar{\lambda}\right) = \lambda, i\left(P, |N|n\bar{\lambda}\right) = i\right\}\right| = \frac{(|N|n\bar{\lambda})! \cdot (|N|n(1-\bar{\lambda})-1)!}{\prod_{j \in N}(n\lambda_j)! \cdot \prod_{j \in N \backslash \{i\}}(n(1-\lambda_j))! \cdot (n(1-\lambda_i)-1)!}. \tag{72}$$

Dividing (72) by (71) yields (43) in a straightforward way.

B Proof of Proposition 6.3

In this appendix, we use the following definitions, notation and properties:

- The function $g : \mathbb{R}_+ \to \mathbb{R}$ is given by

$$g(x) = \begin{cases} x \cdot \ln(x) & \text{if } x > 0, \\ 0 & \text{if } x = 0. \end{cases}$$

- The function $G : [0,1]^N \to \mathbb{R}$ is given by

$$G(\lambda) = |N| \cdot g\left(\bar{\lambda}\right) - \sum_{i \in N} g(\lambda_i) + |N| \cdot g\left(1 - \bar{\lambda}\right) - \sum_{i \in N} g(1 - \lambda_i), \qquad \text{for all } \lambda \in [0,1]^N. \tag{73}$$

- For all $\lambda \in [0,1]^N$, we define

$$N_1^\lambda = \{i \in N : \lambda_i > 0\} \text{ and } N_2^\lambda = \{i \in N : \lambda_i < 1\}. \tag{74}$$

- For $x, y \in \mathbb{R}$ we denote $[x; y]$ as the interval $[\min\{x,y\}, \max\{x,y\}]$, i.e., $[x; y] = [x, y]$ if $x \leq y$ and $[x; y] = [y, x]$ if $x > y$.

- Some arithmetic rules of the Bachmann-Landau notation are given by:

$$\begin{aligned} f(n) = \mathcal{O}(n^a), g(n) = \mathcal{O}(n^b) &\to f(n) + g(n) = \mathcal{O}(n^a), && \text{for all } a \geq b, \\ f(n) = \mathcal{O}(n^a), g(n) = \mathcal{O}(n^{-\infty}) &\to f(n) + g(n) = \mathcal{O}(n^a), && \text{for all } a \in \mathbb{R}, \\ f(n) = \mathcal{O}(n^a), g(n) = \mathcal{O}(n^b) &\to f(n) \cdot g(n) = \mathcal{O}(n^{a+b}), && \text{for all } a, b \in \mathbb{R}, \\ f(n) = \mathcal{O}(n^a) &\to f(n) = \mathcal{O}(n^b), && \text{for all } a \leq b. \end{aligned}$$

  Moreover, we have

$$f(n) = \mathcal{O}(n^a), g(n) = \mathcal{O}^\varepsilon(n^b) \to f(n) + g(n) = \mathcal{O}^\varepsilon(n^a), \qquad \text{for all } a \geq b.$$

- It is well known that for any $k \in \mathbb{R}$, $\delta > 0$ and $c \in (0,1)$ the function $f : \mathbb{N} \to \mathbb{R}_{++}$, defined by $f(n) = n^k \cdot c^{n^\delta}$, is such that $f(n) = \mathcal{O}(n^{-\infty})$.

Lemma B.1 *The function g is continuous and strictly convex, i.e., if $x, y \in \mathbb{R}_+$, $x \neq y$ and $\lambda \in (0,1)$, then $g(\lambda \cdot x + (1-\lambda) \cdot y) < \lambda \cdot g(x) + (1-\lambda) \cdot g(y)$.*

Proof Continuity of f follows from continuity of $x \to x\ln(x)$ for $x > 0$ and the fact that $\lim_{x\downarrow 0} x \cdot \ln(x) = 0$. Strict convexity follows from $g''(x) = \frac{1}{x} > 0$ for every $x > 0$. □

Lemma B.2 *For the function G the following holds:*

1. *G is continuous;*
2. *$G(\lambda) \leq 0$ for all $\lambda \in [0,1]^N$; moreover, $G(\lambda) = 0$ if and only if $\lambda_1 = \lambda_2 = \ldots = \lambda_{|N|}$;*
3. *for all $\lambda \in (0,1)^N$, we have*

$$G(\lambda) = -c(\bar{\lambda}) \cdot \|\lambda - \bar{\lambda} \cdot e_N\|^2 + R,$$

where $|R| \leq \frac{1}{3}|N| \cdot \min\{\lambda_1, \ldots, \lambda_{|N|}, 1-\lambda_1, \ldots, 1-\lambda_{|N|}\}^{-2} \cdot \|\lambda - \bar{\lambda} \cdot e_N\|^3$.

Proof 1. This follows from continuity of g (Lemma B.1).

2. This follows from strict convexity of g (Lemma B.1).

3. Let $\lambda \in (0,1)^N$ and $i \in N$. Then, there exists a $\xi_{i,1} \in [\lambda_i; \bar{\lambda}]$ such that

$$g(\lambda_i) = g(\bar{\lambda}) + g'(\bar{\lambda}) \cdot (\lambda_i - \bar{\lambda}) + \frac{g''(\bar{\lambda})}{2} \cdot (\lambda_i - \bar{\lambda})^2 + \frac{g'''(\xi_{i,1})}{6} \cdot (\lambda_i - \bar{\lambda})^3 \tag{75}$$

$$= g(\bar{\lambda}) + [\ln(\bar{\lambda}) + 1] \cdot (\lambda_i - \bar{\lambda}) + \frac{1}{2\bar{\lambda}} \cdot (\lambda_i - \bar{\lambda})^2 - \frac{1}{6\xi_{i,1}^2} \cdot (\lambda_i - \bar{\lambda})^3, \tag{76}$$

where (75) follows from Taylor's theorem. Note that

$$\sum_{i \in N} (\lambda_i - \bar{\lambda}) = 0. \tag{77}$$

Then, summing the expression (76) of $g(\lambda_i)$ for all $i \in N$ yields

$$\sum_{i \in N} g(\lambda_i) = |N| \cdot g(\bar{\lambda}) + \frac{1}{2\bar{\lambda}} \cdot \|\lambda - \bar{\lambda} \cdot e_N\|^2 - \sum_{i \in N} \frac{1}{6\xi_{i,1}^2} \cdot (\lambda_i - \bar{\lambda})^3.$$

Similarly, we obtain

$$\sum_{i \in N} g(1-\lambda_i) = |N| \cdot g(1-\bar{\lambda}) + \frac{1}{2(1-\bar{\lambda})} \cdot \|\lambda - \bar{\lambda} \cdot e_N\|^2 + \sum_{i \in N} \frac{1}{6\xi_{i,2}^2} \cdot (\lambda_i - \bar{\lambda})^3.$$

where $\xi_{i,2} \in [1-\lambda_i; 1-\bar{\lambda}]$ for all $i \in N$. Now the upperbound of $|R|$ follows from $\xi_{i,1} \geq \min\{\lambda_1, \ldots, \lambda_{|N|}\}$, $\xi_{i,2} \geq \min\{1-\lambda_1, \ldots, 1-\lambda_{|N|}\}$ and $|(\lambda_i - \bar{\lambda})^3| \leq \|\lambda - \bar{\lambda} \cdot e_N\|^3$ for all $i \in N$.

□

Lemma B.3 *Let $d, \varepsilon > 0$. Then, for all $\lambda \in G_\varepsilon \cap D^d$, we have*

$$\min\{\lambda_1, \ldots, \lambda_{|N|}, 1-\lambda_1, \ldots, 1-\lambda_{|N|}\} > \varepsilon - d.$$

Proof Let $\lambda \in G_\varepsilon \cap D^d$. Since $\left|\lambda_i - \bar{\lambda}\right| \leq \left\|\lambda - \bar{\lambda} \cdot e_N\right\| < d$, we obtain $\lambda_i > \bar{\lambda} - d$ and $1-\lambda_i > 1-\bar{\lambda} - d$ for all $i \in N$. Moreover, we have $\varepsilon \leq \bar{\lambda} \leq 1-\varepsilon$. Hence, we obtain $\lambda_i > \varepsilon - d$ and $1-\lambda_i > \varepsilon - d$ for all $i \in N$. This concludes the proof. □

Lemma B.4 *For all* (n,λ) *such that* $n \in \mathbb{N}$ *and* $\lambda \in G^n \backslash \{e_\emptyset, e_N\}$, *we have*

$$t^n(\lambda) = \left(e^{G(\lambda)}\right)^n \cdot (2\pi n)^{\frac{1}{2}(1+|N|-|N_1^\lambda|-|N_2^\lambda|)} \cdot \sqrt{|N|} \cdot \frac{(\bar{\lambda}(1-\bar{\lambda}))^{\frac{1}{2}}}{\prod_{i\in N_1^\lambda}\sqrt{\lambda_i}\prod_{i\in N_2^\lambda}\sqrt{1-\lambda_i}} \cdot \left[1+\mathcal{O}\left(\frac{1}{n\min(\{\lambda_j : j \in N_1^\lambda\}\cup\{1-\lambda_j : j\in N_2^\lambda\})}\right)\right], \tag{78}$$

where N_1^λ *and* N_2^λ *are defined in (74).*

Proof Using (42), we obtain for all (n,λ) such that $n \in \mathbb{N}$ and $\lambda \in G^n\backslash\{e_\emptyset, e_N\}$ that

$$\begin{aligned} t^n(\lambda) &= \frac{\prod_{i\in N}\begin{pmatrix} n \\ n\lambda_i\end{pmatrix}}{\begin{pmatrix}|N|n \\ |N|n\bar{\lambda}\end{pmatrix}} \\ &= \frac{(n!)^{|N|}\cdot(|N|n\bar{\lambda})!\cdot(|N|n(1-\bar{\lambda}))!}{(|N|n)!\cdot\prod_{i\in N}((n\lambda_i)!\cdot(n(1-\lambda_i))!)} \\ &= \frac{(n!)^{|N|}\cdot(|N|n\bar{\lambda})!\cdot(|N|n(1-\bar{\lambda}))!}{(|N|n)!\cdot\prod_{i\in N_1^\lambda}(n\lambda_i)!\cdot\prod_{i\in N_2^\lambda}(n(1-\lambda_i))!}. \end{aligned}$$

Taking the logarithm yields

$$\begin{aligned} \ln(t^n(\lambda)) &= |N|\ln(n!) + \ln((|N|n\bar{\lambda})!) + \ln((|N|n(1-\bar{\lambda}))!) - \ln((|N|n)!) \\ &\quad - \sum_{i\in N_1^\lambda}\ln((n\lambda_i)!) - \sum_{i\in N_2^\lambda}\ln((n(1-\lambda_i))!). \end{aligned} \tag{79}$$

Now, using Stirling's approximation, which is given by

$$\ln(n!) = g(n) - n + \frac{1}{2}\ln(2\pi n) + \mathcal{O}\left(\frac{1}{n}\right), \qquad \text{for all } n \in \mathbb{N},$$

formula (79) can be written as

$$\begin{aligned} \ln(t^n(\lambda)) &= |N|g(n) - |N|n + \frac{1}{2}|N|\ln(2\pi n) + \mathcal{O}\left(\frac{1}{n}\right) \\ &\quad + g(|N|n\bar{\lambda}) - |N|n\bar{\lambda} + \frac{1}{2}\ln(2\pi|N|n\bar{\lambda}) + \mathcal{O}\left(\frac{1}{|N|n\bar{\lambda}}\right) \\ &\quad + g(|N|n(1-\bar{\lambda})) - |N|(n(1-\bar{\lambda})) + \frac{1}{2}\ln(2\pi|N|n(1-\bar{\lambda})) + \mathcal{O}\left(\frac{1}{|N|n(1-\bar{\lambda})}\right) \\ &\quad - \left[g(|N|n) - |N|n + \frac{1}{2}\ln(2\pi|N|n) + \mathcal{O}\left(\frac{1}{|N|n}\right)\right] \\ &\quad - \sum_{i\in N_1^\lambda}\left[g(n\lambda_i) - n\lambda_i + \frac{1}{2}\ln(2\pi n\lambda_i) + \mathcal{O}\left(\frac{1}{n\lambda_i}\right)\right] \\ &\quad - \sum_{i\in N_2^\lambda}\left[g(n(1-\lambda_i)) - n(1-\lambda_i) + \frac{1}{2}\ln(2\pi n(1-\lambda_i)) + \mathcal{O}\left(\frac{1}{n(1-\lambda_i)}\right)\right]. \end{aligned}$$

Now, using that $g(xy) = xg(y) + yg(x)$ for all $x, y \geq 0$, $g(0) = 0$, $\sum_{i\in N_1^\lambda}\lambda_i = |N|\bar{\lambda}$ and $\sum_{i\in N_2^\lambda}(1-\lambda_i) = |N|(1-\bar{\lambda})$, we get

$$\ln(t^n(\lambda)) = |N|g(n) - |N|n + \frac{1}{2}|N|\ln(2\pi n) + \mathcal{O}\left(\frac{1}{n}\right)$$

$$
\begin{aligned}
&+ \bar{\lambda} g(|N|n) + |N|ng(\bar{\lambda}) - |N|n\bar{\lambda} + \frac{1}{2}\ln(2\pi n) + \frac{1}{2}\ln(|N|) + \frac{1}{2}\ln(\bar{\lambda}) + \mathcal{O}\left(\frac{1}{n\bar{\lambda}}\right) \\
&+ (1-\bar{\lambda}) g(|N|n) + |N|ng(1-\bar{\lambda}) - |N|n(1-\bar{\lambda}) + \frac{1}{2}\ln(2\pi n) + \frac{1}{2}\ln(|N|) + \frac{1}{2}\ln(1-\bar{\lambda}) \\
&+ \mathcal{O}\left(\frac{1}{n(1-\bar{\lambda})}\right) \\
&- g(|N|n) + |N|n - \frac{1}{2}\ln(2\pi n) - \frac{1}{2}\ln(|N|) + \mathcal{O}\left(\frac{1}{n}\right) \\
&- |N|\bar{\lambda} g(n) - \sum_{i\in N} ng(\lambda_i) + |N|n\bar{\lambda} - \frac{1}{2}|N_1^\lambda|\ln(2\pi n) - \frac{1}{2}\sum_{i\in N_1^\lambda}\ln(\lambda_i) + \sum_{i\in N_1^\lambda}\mathcal{O}\left(\frac{1}{n\lambda_i}\right) \\
&- |N|(1-\bar{\lambda}) g(n) - \sum_{i\in N} ng(1-\lambda_i) + |N|n(1-\bar{\lambda}) - \frac{1}{2}|N_2^\lambda|\ln(2\pi n) - \frac{1}{2}\sum_{i\in N_2^\lambda}\ln(1-\lambda_i) \\
&+ \sum_{i\in N_2^\lambda}\mathcal{O}\left(\frac{1}{n(1-\lambda_i)}\right).
\end{aligned}
$$

From $|N|g(n) - |N|\bar{\lambda}g(n) - |N|(1-\bar{\lambda})g(n) = 0$, $-|N|n - |N|n\bar{\lambda} - |N|(n(1-\bar{\lambda})) + |N|n + |N|n\bar{\lambda} + |N|n(1-\bar{\lambda}) = 0$, $\bar{\lambda}g(|N|n) + (1-\bar{\lambda})g(|N|n) - g(|N|n) = 0$ and rearranging and collecting some terms follows that

$$
\begin{aligned}
\ln(t^n(\lambda)) = n \cdot &\left[|N|g(\bar{\lambda}) - \sum_{i\in N} g(\lambda_i) + |N|g(1-\bar{\lambda}) - \sum_{i\in N} g(1-\lambda_i)\right] \\
&+ \left[\frac{1}{2}(1 + |N| - |N_1^\lambda| - |N_2^\lambda|)\right]\cdot\ln(2\pi n) + \frac{1}{2}\ln(|N|) \\
&+ \frac{1}{2}\ln(\bar{\lambda}) + \frac{1}{2}\ln(1-\bar{\lambda}) - \frac{1}{2}\sum_{i\in N_1^\lambda}\ln(\lambda_i) - \frac{1}{2}\sum_{i\in N_2^\lambda}\ln(1-\lambda_i) \\
&+ \mathcal{O}\left(\frac{1}{n}\right) + \mathcal{O}\left(\frac{1}{n\bar{\lambda}}\right) + \mathcal{O}\left(\frac{1}{n(1-\bar{\lambda})}\right) + \sum_{i\in N_1^\lambda}\mathcal{O}\left(\frac{1}{n\lambda_i}\right) + \sum_{i\in N_2^\lambda}\mathcal{O}\left(\frac{1}{n(1-\lambda_i)}\right).
\end{aligned}
$$

Then, recall the function G from (73). We get

$$
\begin{aligned}
\ln(t^n(\lambda)) = nG(\lambda) + &\left[\frac{1}{2}(1 + |N| - |N_1^\lambda| - |N_2^\lambda|)\right]\cdot\ln(2\pi n) + \frac{1}{2}\ln(|N|) + \frac{1}{2}\ln(\bar{\lambda}(1-\bar{\lambda})) - \frac{1}{2}\sum_{i\in N_1^\lambda}\ln(\lambda_i) \\
&- \frac{1}{2}\sum_{i\in N_2^\lambda}\ln(1-\lambda_i) + \mathcal{O}\left(\frac{1}{n}\right) + \mathcal{O}\left(\frac{1}{n\bar{\lambda}}\right) + \mathcal{O}\left(\frac{1}{n(1-\bar{\lambda})}\right) + \sum_{i\in N_1^\lambda}\mathcal{O}\left(\frac{1}{n\lambda_i}\right) + \sum_{i\in N_2^\lambda}\mathcal{O}\left(\frac{1}{n(1-\lambda_i)}\right).
\end{aligned}
$$

So, taking the exponent and using the fact that $e^x = 1 + \mathcal{O}(x)$ if $x \in [0, K]$ for some constant $K > 0$, yields

$$
\begin{aligned}
t^n(\lambda) = &\left(e^{G(\lambda)}\right)^n \cdot (2\pi n)^{\frac{1}{2}(1+|N|-|N_1^\lambda|-|N_2^\lambda|)} \cdot \sqrt{|N|} \cdot \frac{(\bar{\lambda}(1-\bar{\lambda}))^{\frac{1}{2}}}{\prod_{i\in N_1^\lambda}\sqrt{\lambda_i}\cdot\prod_{i\in N_2^\lambda}\sqrt{1-\lambda_i}} \\
&\cdot \left[1 + \mathcal{O}\left(\frac{1}{n}\right)\right]\cdot\left[1 + \mathcal{O}\left(\frac{1}{n\bar{\lambda}}\right)\right]\cdot\left[1 + \mathcal{O}\left(\frac{1}{n(1-\bar{\lambda})}\right)\right]\prod_{i\in N_1^\lambda}\left[1 + \mathcal{O}\left(\frac{1}{n\lambda_i}\right)\right]\prod_{i\in N_2^\lambda}\left[1 + \mathcal{O}\left(\frac{1}{n(1-\lambda_i)}\right)\right].
\end{aligned}
$$

Then, as $\lambda_i \geq \min\{\lambda_j : j \in N_1^\lambda\}$ for all $i \in N_1^\lambda$, $1-\lambda_i \geq \min\{1-\lambda_j : j \in N_2^\lambda\}$ for all $i \in N_2^\lambda$, $\bar{\lambda} \geq \frac{1}{|N|}\min\{\lambda_j : j \in N_1^\lambda\}$ and $1-\bar{\lambda} \geq \frac{1}{|N|}\min\{1-\lambda_j : j \in N_2^\lambda\}$, the result follows in a straightforward way. $\square$

Lemma B.5 *We have for all $(\varepsilon, n, \lambda) \in Dom$ with $d_n < \frac{1}{2}\varepsilon$ and $\lambda \in D(n)$ that*

$$\frac{(\bar{\lambda}(1-\bar{\lambda}))^{\frac{1}{2}|N|}}{\prod_{i\in N}\sqrt{\lambda_i}\cdot\prod_{i\in N}\sqrt{1-\lambda_i}} = 1 + \mathcal{O}^{\varepsilon}(n^{-1+\frac{1}{4|N|}}). \tag{80}$$

Proof According to Lemma B.3 we have $\lambda_i \geq \frac{1}{2}\varepsilon$ and $1-\lambda_i \geq \frac{1}{2}\varepsilon$ for all $i \in N$. Consequently, we have $\bar{\lambda} \geq \frac{1}{2}\varepsilon$ and $1-\bar{\lambda} \geq \frac{1}{2}\varepsilon$. According to Taylor's theorem, we have

$$\ln(\lambda_i) = \ln(\bar{\lambda}) + \frac{1}{\bar{\lambda}}\cdot(\lambda_i - \bar{\lambda}) - \frac{1}{2\xi_{i,1}^2}\cdot(\lambda_i-\bar{\lambda})^2, \tag{81}$$

for some $\xi_{i,1} \in [\lambda_i;\bar{\lambda}]$ and for all $i \in N$. From (77) and (81) follows that

$$\frac{1}{2}\sum_{i\in N}\ln(\lambda_i) = \frac{1}{2}|N|\cdot\ln(\bar{\lambda}) - \sum_{i\in N}\frac{1}{4\xi_{i,1}^2}\cdot(\lambda_i-\bar{\lambda})^2. \tag{82}$$

Similarly, we obtain

$$\frac{1}{2}\sum_{i\in N}\ln(1-\lambda_i) = \frac{1}{2}|N|\cdot\ln(1-\bar{\lambda}) - \sum_{i\in N}\frac{1}{4\xi_{i,2}^2}\cdot(\bar{\lambda}-\lambda_i)^2, \tag{83}$$

where $\xi_{i,2} \in [1-\lambda_i; 1-\bar{\lambda}]$ for all $i \in N$. Since $\xi_{i,1} \geq \frac{1}{2}\varepsilon$, $\xi_{i,2} \geq \frac{1}{2}\varepsilon$ and $(\lambda_i - \bar{\lambda})^2 \leq \|\lambda - \bar{\lambda}\cdot e_N\|^2$ for all $i \in N$, we get

$$\begin{aligned}
\sum_{i\in N}\frac{1}{4\xi_{i,1}^2}\cdot(\lambda_i-\bar{\lambda})^2 + \sum_{i\in N}\frac{1}{4\xi_{i,2}^2}\cdot(\bar{\lambda}-\lambda_i)^2 &\leq 2|N|\cdot\varepsilon^{-2}\cdot\|\lambda-\bar{\lambda}\cdot e_N\|^2 \\
&\leq 2|N|\cdot\varepsilon^{-2}\cdot d_n^2 \\
&= 2|N|\cdot\varepsilon^{-2}\cdot n^{-1+\frac{1}{4|N|}} \\
&= \mathcal{O}^{\varepsilon}(n^{-1+\frac{1}{4|N|}}).
\end{aligned}$$

Using the fact that $e^x = 1 + \mathcal{O}(x)$ if $x \in [0, K]$ for some constant $K > 0$ yields

$$e^{\mathcal{O}^{\varepsilon}(n^{-1+\frac{1}{4|N|}})} = 1 + \mathcal{O}^{\varepsilon}(n^{-1+\frac{1}{4|N|}}).$$

Now taking the exponent in (82) and (83) yields the desired result. □

Proof of Proposition 6.3: We prove the result step-by-step: (46) is shown in Lemma B.6, (47) in Lemma B.7, (48) in Lemma B.8 and (49) in Lemma B.9. We implicitly use in the statement of this proposition that if $g(n) = \mathcal{O}(n^c)$ for some $c \leq -\frac{1}{4}$, we have $g(n) = \mathcal{O}(n^{-\frac{1}{4}})$. Note that the result follows directly if $|N| = 1$, so we let $|N| \geq 2$.

Lemma B.6 *We have for all $(\varepsilon, n, \lambda) \in Dom$ that*

$$t^n(\lambda) = \left(e^{-c(\bar{\lambda})\cdot n\cdot\|\lambda-\bar{\lambda}\cdot e_N\|^2}\right)\cdot b(n,\bar{\lambda})\cdot\left[1+\mathcal{O}^{\varepsilon}\left(n^{-\frac{1}{2}+\frac{3}{8|N|}}\right)\right], \qquad \text{if } \lambda \in D(n).$$

Proof It is sufficient to show this result for all $n \in \mathbb{N}$ such that $d_n < \frac{1}{2}\varepsilon$. From Lemma B.3, we then get

$$\min\{\lambda_1,\ldots,\lambda_{|N|}, 1-\lambda_1,\ldots,1-\lambda_{|N|}\} > \frac{1}{2}\varepsilon, \tag{84}$$

and, so,

$$N_1^{\lambda} = N_2^{\lambda} = N. \tag{85}$$

Using Lemma B.2.3 and the fact that $\|\lambda - \bar{\lambda} \cdot e_N\|^3 = \mathcal{O}(n^{-\frac{3}{2}+\frac{3}{8|N|}})$, we get that

$$G(\bar{\lambda}) = -c(\bar{\lambda}) \cdot \|\lambda - \bar{\lambda} \cdot e_N\|^2 + \mathcal{O}^{\varepsilon}(n^{-\frac{3}{2}+\frac{3}{8|N|}}).$$

Hence,

$$\begin{aligned} e^{nG(\lambda)} &= e^{-c(\bar{\lambda})\cdot n \cdot \|\lambda - \bar{\lambda}\cdot e_N\|^2} \cdot e^{\mathcal{O}^{\varepsilon}(n^{-\frac{3}{2}+\frac{3}{8|N|}})} \\ &= e^{-c(\bar{\lambda})\cdot n \cdot \|\lambda - \bar{\lambda}\cdot e_N\|^2} \cdot \left[1 + \mathcal{O}^{\varepsilon}\left(n^{-\frac{1}{2}+\frac{3}{8|N|}}\right)\right]. \end{aligned} \tag{86}$$

where (86) follows from the fact that $e^x = 1 + \mathcal{O}(x)$ if $x \in [0, K]$ for some constant $K > 0$. Substituting (80), (84), (85) and (86) in (78) yields the desired result. $\square$

Lemma B.7 *We have for all $(\varepsilon, n, \lambda) \in Dom$ that*

$$t^n(\lambda) = \mathcal{O}^{\varepsilon}(n^{-\infty}), \qquad \text{if } \lambda \notin D(n).$$

Proof Let $\varepsilon \in (0,1)$, denote $d = \frac{1}{3|N|} \cdot \varepsilon^2$ and recall the function G in (73). The set $G_\varepsilon \backslash D^d$ is compact. Moreover, the function G is continuous (Lemma B.2.1). Hence, the function G takes a maximum value m_ε on $G_\varepsilon \backslash D^d$. As $\lambda \in D^d$ if $\lambda_1 = \ldots = \lambda_{|N|}$, we obtain from Lemma B.2.2 that $m_\varepsilon < 0$. Let (n, λ) be such that $n \in \mathbb{N}$ and $\lambda \in G^n_\varepsilon \backslash D^d$. Since $\lambda_i \geq \frac{1}{n}$ for all $i \in N_1^\lambda$, $1 - \lambda_i \geq \frac{1}{n}$ for all $i \in N_2^\lambda$ and $\bar{\lambda}(1-\bar{\lambda}) < 1$, we get from Lemma B.4 that

$$t^n(\lambda) = \mathcal{O}^{\varepsilon}(n^{\frac{1}{2}(1+|N|)} \cdot (e^{m_\varepsilon})^n).$$

Since $e^{m_\varepsilon} \in (0,1)$ and $\lim_{n\to\infty} c^n \cdot n^d = 0$ for $c \in (0,1)$ and $d \in \mathbb{R}$, we have for all $(\varepsilon, n, \lambda) \in Dom$ that

$$t^n(\lambda) = \mathcal{O}^{\varepsilon}(n^{-\infty}), \qquad \text{if } \lambda \notin D^d.$$

Next, we show this result for all (n, λ) such that $n \in \mathbb{N}$ and $\lambda \in (G^n_\varepsilon \cap D^d) \backslash D(n)$. We obtain from Lemma B.2.3 that

$$\begin{aligned} G(\lambda) &= -c(\bar{\lambda}) \cdot \|\lambda - \bar{\lambda} \cdot e_N\|^2 + R \\ &= -c(\bar{\lambda}) \cdot \|\lambda - \bar{\lambda} \cdot e_N\|^2 \cdot \left[1 - R \cdot (c(\bar{\lambda}))^{-1} \cdot \|\lambda - \bar{\lambda} \cdot e_N\|^{-2}\right], \end{aligned}$$

where $|R| \leq \frac{1}{3}|N| \cdot \min\{\lambda_1, \ldots, \lambda_{|N|}, 1-\lambda_1, \ldots, 1-\lambda_{|N|}\}^{-2} \cdot \|\lambda - \bar{\lambda} \cdot e_N\|^3$. From Lemma B.3, we get

$$\min\{\lambda_1, \ldots, \lambda_{|N|}, 1-\lambda_1, \ldots, 1-\lambda_{|N|}\} > \varepsilon - d > \frac{3|N|-1}{3|N|} \cdot \varepsilon > \frac{1}{2}\varepsilon.$$

Moreover, we have $(c(\bar{\lambda}))^{-1} = 2\bar{\lambda}(1-\bar{\lambda}) \leq \frac{1}{2}$ and $\|\lambda - \bar{\lambda} \cdot e_N\| < d$. Therefore, we have

$$|R \cdot (c(\bar{\lambda}))^{-1} \cdot \|\lambda - \bar{\lambda} \cdot e_N\|^{-2}| \leq |R| \cdot (c(\bar{\lambda}))^{-1} \cdot \|\lambda - \bar{\lambda} \cdot e_N\|^{-2} \leq \frac{1}{6}|N| \cdot \left(\frac{1}{2}\varepsilon\right)^{-2} \cdot d < \frac{1}{2}.$$

So, then, we obtain that

$$nG(\lambda) < -c(\bar{\lambda}) \cdot n \cdot \|\lambda - \bar{\lambda} \cdot e_N\|^2 \cdot \frac{1}{2} \leq -n \cdot \|\lambda - \bar{\lambda} \cdot e_N\|^2 \leq -n^{\frac{1}{4|N|}},$$

which follows from $c(\bar{\lambda}) \geq 2$, and, hence,

$$e^{nG(\lambda)} < e^{-n^{\frac{1}{4|N|}}}. \tag{87}$$

We get

$$t^n(\lambda) = \mathcal{O}^{\varepsilon}(e^{nG(\lambda)} \cdot n^{\frac{1}{2}(1-|N|)}) \tag{88}$$

$$= \mathcal{O}^\varepsilon((e^{-1})^{n^{\frac{1}{4|N|}}} \cdot n^{\frac{1}{2}(1-|N|)}) \tag{89}$$
$$= \mathcal{O}^\varepsilon(n^{-\infty}), \tag{90}$$

where (88) follows from Lemma B.4, (89) follows from (87) and (90) follows from the fact that $\lim_{n\to\infty} n^k \cdot c^{n^\delta} = 0$ for all $k \in \mathbb{R}$, $c \in (0,1)$ and $\delta > 0$. □

Lemma B.8 *We have for all $i \in N$ and $(\varepsilon, n, \lambda) \in Dom$ that*

$$p_i^n(\lambda) = \frac{1}{|N|} \cdot \left[1 + \mathcal{O}^\varepsilon\left(n^{-\frac{1}{2}+\frac{1}{8|N|}}\right)\right], \qquad \textit{if } \lambda \in D(n).$$

Proof Note that from $\lambda \in G_\varepsilon^n$ it follows that $\lambda \neq e_N$, so $\bar{\lambda} < 1$. Then, the result follows directly from

$$\begin{aligned}
\left|\frac{1-\lambda_i}{\sum_{j\in N}(1-\lambda_j)} - \frac{1}{|N|}\right| &= \left|\frac{1-\lambda_i}{(1-\bar{\lambda})|N|} - \frac{1-\bar{\lambda}}{(1-\bar{\lambda})|N|}\right| \\
&= \frac{|\bar{\lambda}-\lambda_i|}{(1-\bar{\lambda})|N|} \\
&< \frac{n^{-\frac{1}{2}+\frac{1}{8|N|}}}{(1-\bar{\lambda})|N|} \qquad (91) \\
&\leq \frac{n^{-\frac{1}{2}+\frac{1}{8|N|}}}{\varepsilon|N|}, \qquad (92)
\end{aligned}$$

for all $(\varepsilon, n, \lambda) \in Dom$ such that $\lambda \in D(n)$. Here, (91) follows from $|\bar{\lambda} - \lambda_i| \leq \|\lambda - \bar{\lambda} \cdot e_N\| < d_n = n^{-\frac{1}{2}+\frac{1}{8|N|}}$ and (92) follows from $1 - \bar{\lambda} \geq \varepsilon$. This concludes the proof. □

Lemma B.9 *We have for all $i \in N$ and $(\varepsilon, n, \lambda) \in Dom$ that*

$$p_i^n(\lambda) = \mathcal{O}(1).$$

Proof This follows directly from $0 \leq p_i^n(\lambda) \leq 1$. □

C Proof of Proposition 6.5

In this appendix, we use the following notation:

- For all $x \in \mathbb{R}$, we write $\lfloor x \rfloor$ as the largest integer not greater than x and $\lceil x \rceil$ as the smallest integer not less than x.
- For all $n \in \mathbb{N}$ and $\lambda \in G^n$, the set $C^n(\lambda)$ is given by

$$C^n(\lambda) = \left\{\lambda + \frac{1}{n} \cdot x : x \in [0,1]^N\right\}. \tag{93}$$

- The set $D'(n)$ is given by

$$D'(n) = D^{d_n'}, \qquad \text{where } d_n' = d_n + (\sqrt{|N|}/n) = n^{-\frac{1}{2}+\frac{1}{8|N|}} + \frac{\sqrt{|N|}}{n}. \tag{94}$$

- If there might be confusion about the notation $|\cdot|$ for the absolute value of a real number and the cardinality of a set, we sometimes write $\sharp(A)$ as the cardinality of the set A.

- We write $\nu(B)$ as the Lebesgue measure of the set B. Note that

$$\nu(C^n(\lambda)) = n^{-|N|}, \qquad \text{for all } \lambda \in G^n, \tag{95}$$

and

$$\nu(D'(n)) = \mathcal{O}(d_n'^{|N|-1}) = \mathcal{O}(n^{-\frac{1}{2}+\frac{1}{8|N|}}), \qquad \text{for all } n \in \mathbb{N}. \tag{96}$$

- Let $R \in \mathcal{R}$ and $\varepsilon > 0$. We define the set $B(R,n)$ by

$$B(R,n) = \left\{ \lambda \in [0,1]^N : \exists \hat{\lambda} \in [0,1]^N \backslash L(R) : \|\lambda - \hat{\lambda}\| < \frac{1}{n} \right\}, \tag{97}$$

for all $R \in \mathcal{R}$ and $n \in \mathbb{N}$, where $L(R)$ is defined in Definition 4.5. This is the set of all participation profiles close to a participation profile that is an element of multiple sets K_m. As the risk capital allocation problem is always clear from the context, we write $B(n) = B(R,n)$.

First, we show that only the participation profiles in G^n_ε have a non-negligible aggregate contribution.

Lemma C.1 *For all $i \in N$, we have*

$$n^{-1} \cdot \sum_{\lambda \in G^n \backslash G^n_\varepsilon : \lambda_i < 1} t^n(\lambda) \cdot p^n_i(\lambda) = \mathcal{O}(\varepsilon) + \mathcal{O}(n^{-1}).$$

Proof Recall (59) for the definition of $\widetilde{G}^n_k$. Then, we obtain

$$\begin{aligned}
\sum_{\lambda \in G^n \backslash G^n_\varepsilon : \lambda_i < 1} t^n(\lambda) \cdot p^n_i(\lambda) &= \sum_{k=0}^{\lceil \varepsilon |N| n \rceil - 1} \sum_{\lambda \in \widetilde{G}^n_k : \lambda_i < 1} t^n(\lambda) \cdot p^n_i(\lambda) + \sum_{k=\lfloor (1-\varepsilon)|N|n \rfloor + 1}^{|N|n-1} \sum_{\lambda \in \widetilde{G}^n_k : \lambda_i < 1} t^n(\lambda) \cdot p^n_i(\lambda) && (98)\\
&\leq \sum_{k=0}^{\lceil \varepsilon |N| n \rceil - 1} \sum_{\lambda \in \widetilde{G}^n_k : \lambda_i < 1} t^n(\lambda) + \sum_{k=\lfloor (1-\varepsilon)|N|n \rfloor + 1}^{|N|n-1} \sum_{\lambda \in \widetilde{G}^n_k : \lambda_i < 1} t^n(\lambda) && (99)\\
&\leq \sum_{k=0}^{\lceil \varepsilon |N| n \rceil - 1} 1 + \sum_{k=\lfloor (1-\varepsilon)|N|n \rfloor + 1}^{|N|n-1} 1 && (100)\\
&= \lceil \varepsilon |N| n \rceil + \lceil \varepsilon |N| n \rceil - 1 \\
&< 2\varepsilon |N| n + 1 && (101)\\
&= \mathcal{O}(\varepsilon) \cdot n + \mathcal{O}(1).
\end{aligned}$$

Here, (98) follows from (59) and (68), (99) follows from $0 \leq p^n_i(\lambda) \leq 1$ for all $\lambda \in G^n \backslash \{e_N\}$, (100) follows from $\sum_{\lambda \in \widetilde{G}^n_k} t^n(\lambda) = 1$ for all $k \in \{0, \ldots, |N|n - 1\}$ and (101 follows from the fact that $\lceil x \rceil < x + 1$ for all $x \in \mathbb{R}$. $\square$

The following result follows almost directly from Proposition 6.3.

Lemma C.2 *For all $i \in N$, we have*

$$\sum_{\lambda \in [G^n_\varepsilon \backslash D(n)] : \lambda_i < 1} t^n(\lambda) \cdot p^n_i(\lambda) = \mathcal{O}^\varepsilon(n^{-\infty}).$$

Proof This result follows directly from

$$\begin{aligned}
\sum_{\lambda \in [G^n_\varepsilon \backslash D(n)] : \lambda_i < 1} t^n(\lambda) \cdot p^n_i(\lambda) &= \sum_{\lambda \in [G^n_\varepsilon \backslash D(n)] : \lambda_i < 1} \mathcal{O}^\varepsilon(n^{-\infty}) && (102)\\
&< (n+1)^{|N|} \cdot \mathcal{O}^\varepsilon(n^{-\infty}) && (103)\\
&= \mathcal{O}^\varepsilon(n^{-\infty}),
\end{aligned}$$

where (102) follows from Proposition 6.3 and (103) follows from $\sharp(\{\lambda \in [G^n_\varepsilon \backslash D(n)] : \lambda_i < 1\}) < \sharp(G^n) = (n+1)^{|N|}$.
$\square$

Lemma C.3 *Let $R \in \mathcal{R}$. Then, we have*

$$r\left(\lambda + (1/n) \cdot e_i\right) - r(\lambda) = \mathcal{O}\left(n^{-1}\right),$$

for all $i \in N$ and (n, λ) such that $n \in \mathbb{N}$, $\lambda \in G^n$ and $\lambda_i < 1$.

Proof Denote $c = \max\{|f_{\mathbb{Q}}(e_j)| : \mathbb{Q} \in Q(\rho), j \in N\}$. Let $\mathbb{Q}_1, \mathbb{Q}_2 \in Q(\rho)$ be such that $r\left(\lambda + (1/n) \cdot e_i\right) = f_{\mathbb{Q}_1}\left(\lambda + (1/n) \cdot e_i\right)$ and $r(\lambda) = f_{\mathbb{Q}_2}(\lambda)$. Then, we have

$$\begin{aligned} r\left(\lambda + (1/n) \cdot e_i\right) - r(\lambda) &= f_{\mathbb{Q}_1}\left(\lambda - (1/n) \cdot e_i\right) - f_{\mathbb{Q}_2}(\lambda) \\ &\leq f_{\mathbb{Q}_1}\left(\lambda + (1/n) \cdot e_i\right) - f_{\mathbb{Q}_1}(\lambda) \\ &= \frac{1}{n} \cdot f_{\mathbb{Q}_1}(e_i) \\ &\leq \frac{1}{n} \cdot c \end{aligned}$$

and

$$\begin{aligned} r\left(\lambda + (1/n) \cdot e_i\right) - r(\lambda) &= f_{\mathbb{Q}_1}\left(\lambda + (1/n) \cdot e_i\right) - f_{\mathbb{Q}_2}(\lambda) \\ &\geq f_{\mathbb{Q}_2}\left(\lambda + (1/n) \cdot e_i\right) - f_{\mathbb{Q}_2}(\lambda) \\ &= \frac{1}{n} \cdot f_{\mathbb{Q}_2}(e_i) \\ &\geq -\frac{1}{n} \cdot c. \end{aligned}$$

This concludes the proof. □

Lemma C.4 *For all $i \in N$, we have*

$$\sum_{\lambda \in G_\varepsilon^n \cap D(n)} |t^n(\lambda) \cdot p_i^n(\lambda) - h^n(\lambda)| = \mathcal{O}^\varepsilon(n^{\frac{3}{4}}).$$

Proof It is sufficient to show this result only for $n \in \mathbb{N}$ such that $d_n < \frac{1}{2}\varepsilon$. If $|N| = 1$ the result is trivial as $t^n(\lambda) \cdot p_i^n(\lambda) = h^n(\lambda) = 1$ for all $\lambda \in G_\varepsilon^n$. Next, we let $|N| \geq 2$. For all $\lambda \in G_\varepsilon^n \cap D(n)$, we have

$$\begin{aligned} |t^n(\lambda) \cdot p_i^n(\lambda) - h^n(\lambda)| &= \left| h^n(\lambda) \cdot [1 + \mathcal{O}^\varepsilon(n^{-\frac{1}{2} + \frac{3}{8|N|}})] \cdot [1 + \mathcal{O}^\varepsilon(n^{-\frac{1}{2} + \frac{1}{8|N|}})] - h^n(\lambda) \right| \quad (104) \\ &= \left| h^n(\lambda) \cdot \mathcal{O}^\varepsilon(n^{-\frac{1}{2} + \frac{3}{8|N|}}) \right| \\ &= \mathcal{O}^\varepsilon(n^{-\frac{1}{2}|N| + \frac{3}{8|N|}}), \quad (105) \end{aligned}$$

where (104) follows from Lemma B.6 and Lemma B.8 and (105) follows from $h^n(\lambda) = \mathcal{O}^\varepsilon(n^{\frac{1}{2}(1-|N|)})$. If $y \in C^n(\lambda)$ for a $\lambda \in G_\varepsilon^n \cap D(n)$, we have

$$\begin{aligned} \|y - \bar{y} \cdot e_N\| &\leq \|y - \bar{\lambda} \cdot e_N\| \quad (106) \\ &\leq \|y - \lambda\| + \|\lambda - \bar{\lambda} \cdot e_N\| \quad (107) \\ &< (\sqrt{|N|}/n) + n^{-\frac{1}{2} + \frac{1}{8|N|}}, \quad (108) \end{aligned}$$

where (106) and (107) follow from the triangular inequality and (108) follows from the fact that $\|y - \lambda\| \leq (\sqrt{|N|}/n)$ for all $y \in C^n(\lambda)$. So, we get

$$\bigcup_{\lambda \in G_\varepsilon^n \cap D(n)} C^n(\lambda) \subset D'(n). \quad (109)$$

From this and (95), we get

$$n^{-|N|} \cdot \sharp(G_\varepsilon^n \cap D(n)) \leq \nu\left(D'(n)\right) \quad (110)$$

$$= \mathcal{O}(d_n'^{|N|-1}) \tag{111}$$
$$= \mathcal{O}\left(\left(n^{-\frac{1}{2}+\frac{1}{8|N|}} + (\sqrt{|N|}/n)\right)^{(|N|-1)}\right)$$
$$= \mathcal{O}(n^{-\frac{1}{2}|N|+\frac{5}{8}-\frac{1}{8|N|}}), \tag{112}$$

where (110) follows from (95) and (109), and (111) follows from (96). From this, we get

$$\sum_{\lambda \in G_\varepsilon^n \cap D(n)} |t^n(\lambda) \cdot p_i^n(\lambda) - h^n(\lambda)| \leq \sharp(G_\varepsilon^n \cap D(n)) \cdot \mathcal{O}^\varepsilon(n^{-\frac{1}{2}|N|+\frac{3}{8|N|}})$$
$$= \mathcal{O}^\varepsilon(n^{\frac{5}{8}+\frac{2}{8|N|}}).$$

As $|N| \geq 2$, this concludes the proof. □

Lemma C.5 *Let $R \in \mathcal{R}$. Then, for all $\varepsilon > 0$ and all $m \in \{p^*+1, \ldots, p\}$, we have for sufficiently large n that*

$$G_\varepsilon \cap D(n) \cap K_m = \emptyset.$$

Proof If $p^* = p$, the result follows directly and, so, we let $p^* < p$. Denote

$$\alpha = r(e_N) - \max_{m' \in \{p^*+1,\ldots,p\}} f_{\mathbb{Q}_{m'}}(e_N) > 0,$$

and let $\ell \in \{1, \ldots, p^*\}$ and $m \in \{p^*+1, \ldots, p\}$. Then, we have

$$f_{\mathbb{Q}_\ell}(e_N) \geq f_{\mathbb{Q}_m}(e_N) + \alpha.$$

By linearity of $f_{\mathbb{Q}_\ell}$, we have

$$f_{\mathbb{Q}_\ell}(t \cdot e_N) - f_{\mathbb{Q}_m}(t \cdot e_N) = t \cdot (f_{\mathbb{Q}_\ell}(e_N) - f_{\mathbb{Q}_m}(e_N)) \geq t \cdot \alpha, \qquad \text{for all } t \in [0,1]. \tag{113}$$

If $f_{\mathbb{Q}_{m'}}(e_i) = 0$ for all $m' \in \{1, \ldots, p\}$ and for all $i \in N$, we have $p = p^* = 1$, which contradicts the assumption that $p^* < p$. So, let $M = \max_{m' \in \{1,\ldots,p\}} \|(f_{\mathbb{Q}_{m'}}(e_i))_{i \in N}\| > 0$ and $\varepsilon > 0$. Then, define $N_\varepsilon = \left(\frac{2 \cdot M}{\alpha \cdot \varepsilon}\right)^4$ and let $n > N_\varepsilon$. Then, we obtain for every $\lambda \in G_\varepsilon \cap D(n)$ that

$$f_{\mathbb{Q}_\ell}(\lambda) - f_{\mathbb{Q}_m}(\lambda) = f_{\mathbb{Q}_\ell}(\bar\lambda \cdot e_N) - f_{\mathbb{Q}_m}(\bar\lambda \cdot e_N) + f_{\mathbb{Q}_\ell}(\lambda - \bar\lambda \cdot e_N) - f_{\mathbb{Q}_m}(\lambda - \bar\lambda \cdot e_N) \tag{114}$$
$$\geq \bar\lambda\alpha + f_{\mathbb{Q}_\ell}(\lambda - \bar\lambda \cdot e_N) - f_{\mathbb{Q}_m}(\lambda - \bar\lambda \cdot e_N), \tag{115}$$

where (114) follows from linearity of $f_{\mathbb{Q}_\ell}$ and $f_{\mathbb{Q}_m}$ and (115) follows from (113). Moreover, we obtain that

$$|f_{\mathbb{Q}_{m'}}(\lambda - \bar\lambda \cdot e_N)| \leq \|(f_{\mathbb{Q}_{m'}}(e_i))_{i \in N}\| \cdot \|\lambda - \bar\lambda \cdot e_N\| \tag{116}$$
$$\leq M \cdot n^{-\frac{1}{4}} \tag{117}$$
$$< M \cdot N_\varepsilon^{-\frac{1}{4}} \tag{118}$$
$$= \frac{1}{2}\varepsilon \cdot \alpha \tag{119}$$
$$\leq \frac{1}{2}\bar\lambda \cdot \alpha, \tag{120}$$

for all $m' \in \{1, \ldots, p\}$, where (116) follows from the Cauchy-Schwartz inequality applied to $\sum_{i \in N} f_{\mathbb{Q}_m}(e_i) \cdot (\lambda_i - \bar\lambda)$, (117) follows from $m' \in \{1, \ldots, p\}$ and $\lambda \in D(n)$, (118) follows from $n > N_\varepsilon$, (119) follows from substituting the definition of N_ε, follows from and (120) follows from $\lambda \in G_\varepsilon$. Hence, substituting (120) in (115) yields that $f_{\mathbb{Q}_\ell}(\lambda) - f_{\mathbb{Q}_m}(\lambda) > 0$. Therefore, we have $\lambda \notin K_m$ for every $\lambda \in G_\varepsilon \cap D(n)$ and, hence,

$$G_\varepsilon \cap D(n) \cap K_m = \emptyset. \tag{121}$$

□

Note that from (23) and Lemma C.5 it follows for all $\varepsilon > 0$ that

$$G_\varepsilon \cap D(n) \subset \bigcup_{m\in\{1,\ldots,p^*\}} K_m, \qquad \text{for large } n.$$

We next show that we can neglect participation profiles close to profiles where the function r is non-differentiable. Note that $B(n)$, as defined in (97), is the set of participation profiles close to a participation profile where the function r is non-differentiable. For all $n \in \mathbb{N}$ we have that if $\lambda \in K_m\backslash B(n)$ for some $m \in \{1,\ldots,p\}$, then $\lambda+(1/n)\cdot e_i \in K_m$ for all $i \in N$ and, by linearity of $f_{\mathbb{Q}_m}$, $r\left(\lambda+(1/n)\cdot e_i\right)-r(\lambda) = \frac{1}{n}\cdot E_{\mathbb{Q}_m}[X_i]$.

Lemma C.6 *Let $R \in \mathcal{R}$. Then, we have*

$$\sum_{\lambda\in G_\varepsilon^n\cap D(n)\cap B(n)} h^n(\lambda) = \mathcal{O}^\varepsilon(n^{\frac{5}{8}}).$$

Proof If $p = 1$, we have that $B(n) = \emptyset$ for all $n \in \mathbb{N}$ and, so, the result follows directly. Next, let $p > 1$. Recall (26), i.e.,

$$[0,1]^N\backslash L(R) = \bigcup_{\ell,m\in\{1,\ldots,p\}:\ell\neq m} K_\ell \cap K_m.$$

Let $\varepsilon > 0$, $\ell, m \in \{1,\ldots,p\}$, $\ell \neq m$ and $n > \frac{2}{\varepsilon}$. We define

$$H^n(\ell,m) = \left\{\lambda \in G_\varepsilon^n \cap D(n) : \exists\hat{\lambda} \in K_\ell \cap K_m : \|\lambda-\hat{\lambda}\| \leq \frac{1}{n}\right\},$$

and $D_\varepsilon = \{\lambda \in G_\varepsilon : \lambda = \bar{\lambda}\cdot e_N\}$. According to Lemma C.5 we have for all $m \in \{p^*+1,\ldots,p\}$ that $D_\varepsilon \cap K_m = \emptyset$. Since D_ε and K_m are both compact we can define $\alpha_{\varepsilon,m} = \operatorname{dist}(D_\varepsilon, K_m) = \min\{\|x-y\| : x \in D_\varepsilon, y \in K_m\}$. Obviously, $\alpha_{\varepsilon,m} > 0$. So, if $\ell \notin \{1,\ldots,p^*\}$ or $m \notin \{1,\ldots,p^*\}$ we get $H^n(\ell,m) = \emptyset$ for large n. If $p^* = 1$ it follows from this that $H^n(\ell,m) = \emptyset$ for all $\ell, m \in \{1,\ldots,p\}$. Next, let $p^* > 1$ and $\ell, m \in \{1,\ldots,p^*\}$. Recall (25) from Lemma 4.6, i.e.,

$$K_\ell \cap K_m \subset \left\{\lambda \in \mathbb{R}^N : \sum_{i\in N} \lambda_i\cdot\left(E_{\mathbb{Q}_\ell}[X_i]-E_{\mathbb{Q}_m}[X_i]\right) = 0\right\} := V(\ell,m).$$

Note that $V(\ell,m)$ is an $(|N|-1)$-dimensional linear space where $\{t\cdot e_N : t \in \mathbb{R}\} \subset V(\ell,m)$. To obtain an upperbound of the cardinality of $H^n(\ell,m)$, we first derive the Lebesgue measure of the following Euclidean set

$$\widetilde{H}^n(\ell,m) = \left\{\lambda \in G_{\frac{1}{2}\varepsilon} \cap D'(n) : \exists\hat{\lambda} \in V(\ell,m) : \|\lambda-\hat{\lambda}\| \leq \frac{\sqrt{|N|}}{n}+\frac{1}{n}\right\}.$$

We describe this set via the Gram-Schmidt process. Choose an orthonormal basis $u_1,\ldots,u_{|N|}$ of $\mathbb{R}^N$ such that $u_1 = \frac{e_N}{\sqrt{|N|}}$, $u_1,\ldots,u_{|N|-1}$ is an orthonormal basis of the $(|N|-1)$-dimensional space $V(\ell,m)$ and $u_{|N|}$ is a unit normal vector of the $(|N|-1)$-dimensional space $V(\ell,m)$. So $u_{|N|}$ is a multiple of the vector $(E_{\mathbb{Q}_\ell}[X_i]-E_{\mathbb{Q}_m}[X_i])_{i\in N}$. Now let $\lambda \in \widetilde{H}^n(\ell,m)$. Let λ_1 be the unique element in $V(\ell,m)$ that is closest to λ. Obviously $\|\lambda-\lambda_1\| \leq \frac{\sqrt{|N|}}{n}+\frac{1}{n}$. Let $\lambda_2 = \bar{\lambda}_1\cdot e_N$ $(= \bar{\lambda}\cdot e_N)$ be the unique element in $\{t\cdot e_N : t \in \mathbb{R}\}$ that is closest to λ_1 (and hence closest to λ). We provide an overview of the construction of λ_1 and λ_2 in Figure 6. Obviously $\|\lambda-\lambda_2\|^2 = \|\lambda-\lambda_1\|^2+\|\lambda_1-\lambda_2\|^2$ and hence $\|\lambda_1-\lambda_2\| \leq \|\lambda-\lambda_2\| = \|\lambda-\bar{\lambda}\cdot e_N\| < d_n'$. Now we can write $\lambda = \alpha_1\cdot u_1+\ldots+\alpha_{|N|}\cdot u_{|N|}$ where $\lambda_2 = \alpha_1\cdot u_1$, $\lambda_1-\lambda_2 = \alpha_2\cdot u_2+\ldots+\alpha_{|N|-1}\cdot u_{|N|-1}$ and $\lambda-\lambda_1 = \alpha_{|N|}\cdot u_{|N|}$. From this follows that $|\alpha_1| = \|\lambda_2\| = \bar{\lambda}\cdot\sqrt{|N|} < \sqrt{|N|}$, $|\alpha_k| \leq \sqrt{\alpha_2^2+\ldots+\alpha_{|N|-1}^2} = \|\lambda_1-\lambda_2\| < d_n' = \mathcal{O}(n^{-\frac{1}{2}+\frac{1}{8|N|}})$ for all $k \in \{2,\ldots,|N|-1\}$ and $|\alpha_{|N|}| = \|\lambda-\lambda_1\| \leq \frac{1}{n}+\frac{\sqrt{|N|}}{n} = \mathcal{O}(n^{-1})$. Hence,

$$\nu(\widetilde{H}^n(\ell,m)) = \mathcal{O}(1)\cdot\mathcal{O}(n^{(-\frac{1}{2}+\frac{1}{8|N|})(|N|-2)})\cdot\mathcal{O}(n^{-1})$$

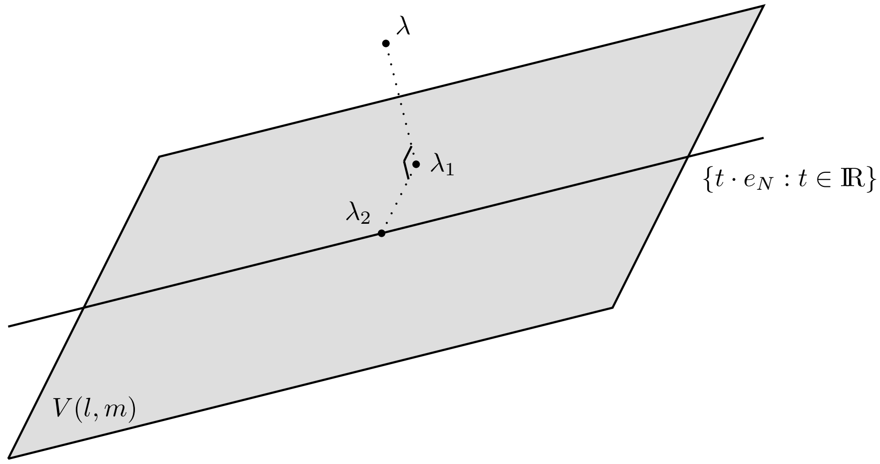


Figure 6: Illustration of λ_1 and λ_2 corresponding to the proof of Lemma C.6.

$$= \mathcal{O}(n^{-\frac{1}{2}|N|+\frac{1}{8}}). \tag{122}$$

For all $\lambda \in G^n_\varepsilon$ and $y \in C^n(\lambda)$, we get from

$$\bar{y} = \bar{\lambda} + (\bar{y} - \bar{\lambda}) \begin{cases} \geq \varepsilon - (1/n) > \frac{1}{2}\varepsilon, \\ \leq 1 - \varepsilon + (1/n) < 1 - \frac{1}{2}\varepsilon, \end{cases} \tag{123}$$

that $y \in G_{\frac{1}{2}\varepsilon}$. Moreover, we get

$$\min_{\hat{\lambda} \in V(\ell,m)} \|y - \hat{\lambda}\| \leq \|y - \lambda\| + \min_{\hat{\lambda} \in V(\ell,m)} \|\lambda - \hat{\lambda}\| \leq \frac{\sqrt{|N|}}{n} + \frac{1}{n}, \qquad \text{for all } \lambda \in H^n(\ell,m) \text{ and } y \in C^n(\lambda).$$

From this, (109) and (123), we get

$$\bigcup_{\lambda \in H^n(\ell,m)} C^n(\lambda) \subset \widetilde{H}^n(\ell,m), \qquad \text{for all } n \in \mathbb{N} \text{ such that } n > \frac{2}{\varepsilon}. \tag{124}$$

From (95) and (124) we get

$$n^{-|N|} \cdot \sharp(H^n(\ell,m)) \leq \nu(\widetilde{H}^n(\ell,m)). \tag{125}$$

Substituting (122) in (125) yields

$$\sharp(H^n(\ell,m)) = \mathcal{O}(n^{\frac{1}{2}|N|+\frac{1}{8}}). \tag{126}$$

Then, we obtain

$$\begin{aligned} \sum_{\lambda \in G^n_\varepsilon \cap D(n) \cap B(n)} h^n(\lambda) &\leq \sum_{\ell,m \in \{1,\ldots,p\}:\ell \neq m} \ \sum_{\lambda \in H^n(\ell,m)} h^n(\lambda) && (127) \\ &\leq \begin{pmatrix} p \\ 2 \end{pmatrix} \cdot \max_{\ell,m \in \{1,\ldots,p\}:\ell \neq m} \sharp(H^n(\ell,m)) \cdot \max_{\lambda \in G_\varepsilon} h^n(\lambda) && (128) \\ &= \begin{pmatrix} p \\ 2 \end{pmatrix} \cdot \mathcal{O}(n^{\frac{1}{2}|N|+\frac{1}{8}}) \cdot \mathcal{O}^\varepsilon(n^{\frac{1}{2}(1-|N|)}) && (129) \\ &= \mathcal{O}^\varepsilon(n^{\frac{5}{8}}), \end{aligned}$$

where (127) follows from (26), (128) follows from $\sharp(\{\ell, m \in \{1,\ldots,p\} : \ell \neq m\}) = \begin{pmatrix} p \\ 2 \end{pmatrix}$ and (129) follows from (126) and $h^n(\lambda) = \mathcal{O}^\varepsilon(n^{\frac{1}{2}(1-|N|)})$ for all $\lambda \in G_\varepsilon$. This concludes the proof. □

Proof of Proposition 6.5: It is sufficient to show this result for sufficiently large n. We get

$$K_i^n(R) = \sum_{\lambda\in G^n:\lambda_i<1} t^n(\lambda)\cdot p_i^n(\lambda)\cdot [r(\lambda+(1/n)\cdot e_i)-r(\lambda)] \tag{130}$$

$$= \sum_{\lambda\in G_\varepsilon^n:\lambda_i<1} t^n(\lambda)\cdot p_i^n(\lambda)\cdot [r(\lambda+(1/n)\cdot e_i)-r(\lambda)] + \mathcal{O}(\varepsilon)+\mathcal{O}(n^{-1}) \tag{131}$$

$$= \sum_{\lambda\in G_\varepsilon^n\cap D(n)} t^n(\lambda)\cdot p_i^n(\lambda)\cdot [r(\lambda+(1/n)\cdot e_i)-r(\lambda)] + \mathcal{O}(\varepsilon)+\mathcal{O}^\varepsilon\left(n^{-1}\right) \tag{132}$$

$$= \sum_{\lambda\in G_\varepsilon^n\cap D(n)} h^n(\lambda)\cdot [r(\lambda+(1/n)\cdot e_i)-r(\lambda)] + \mathcal{O}(\varepsilon)+\mathcal{O}^\varepsilon(n^{-\frac{1}{4}}) \tag{133}$$

$$= \sum_{\lambda\in [G_\varepsilon^n\cap D(n)]\backslash B(n)} h^n(\lambda)\cdot [r(\lambda+(1/n)\cdot e_i)-r(\lambda)] + \mathcal{O}(\varepsilon)+\mathcal{O}^\varepsilon(n^{-\frac{1}{4}}) \tag{134}$$

$$= \sum_{m=1}^{p}\sum_{\lambda\in [G_\varepsilon^n\cap D(n)\cap K_m]\backslash B(n)} h^n(\lambda)\cdot\frac{1}{n}\cdot E_{\mathbb{Q}_m}[X_i] + \mathcal{O}(\varepsilon)+\mathcal{O}^\varepsilon(n^{-\frac{1}{4}}) \tag{135}$$

$$= \sum_{m=1}^{p^*}\sum_{\lambda\in [G_\varepsilon^n\cap D(n)\cap K_m]\backslash B(n)} h^n(\lambda)\cdot\frac{1}{n}\cdot E_{\mathbb{Q}_m}[X_i] + \mathcal{O}(\varepsilon)+\mathcal{O}^\varepsilon(n^{-\frac{1}{4}}) \tag{136}$$

$$= \sum_{m=1}^{p^*} E_{\mathbb{Q}_m}[X_i]\cdot\frac{1}{n}\cdot\sum_{\lambda\in [G_\varepsilon^n\cap D(n)\cap K_m]\backslash B(n)} h^n(\lambda) + \mathcal{O}(\varepsilon)+\mathcal{O}^\varepsilon(n^{-\frac{1}{4}})$$

$$= \sum_{m=1}^{p^*} E_{\mathbb{Q}_m}[X_i]\cdot\frac{1}{n}\cdot\sum_{\lambda\in G_\varepsilon^n\cap D(n)\cap K_m} h^n(\lambda) + \mathcal{O}(\varepsilon)+\mathcal{O}^\varepsilon(n^{-\frac{1}{4}}), \tag{137}$$

where (130) follows from Proposition 6.2, (131) follows from Lemma C.1 and Lemma C.3, (132) follows from Lemma C.2 and Lemma C.3, (133) follows from Lemma C.3 and Lemma C.4, (134) follows from Lemma C.3 and Lemma C.6, (135) follows from $[0,1]^N\backslash L(R)\subset B(n)$, (136) follows from Lemma C.5 and (137) follows from Lemma C.6. This concludes the proof.

D Proof of Proposition 6.6

In this appendix, we use the same notation as in Appendix C.

Lemma D.1 *The function h^n is differentiable for a fixed $n\in\mathbb{N}$, and, moreover, we have for all $i\in N$ and $(\varepsilon,n,\lambda)\in Dom$ that*

$$\frac{\partial h^n}{\partial\lambda_i}(\lambda) = \mathcal{O}^\varepsilon(n^{-\frac{1}{2}|N|+1+\frac{1}{8|N|}}), \qquad \text{if } \lambda\in D'(n),$$

where $D'(n)$ is defined in (94).

Proof Define the functions $f^n(\lambda) = -c(\bar\lambda)\cdot n\cdot\|\lambda-\bar\lambda\cdot e_N\|^2$ and $g(\lambda) = (\bar\lambda(1-\bar\lambda))^{\frac{1}{2}(1-|N|)}$ for all $\lambda\in[0,1]^N$. Then, we obtain

$$\frac{\partial h^n}{\partial\lambda_i}(\lambda) = \frac{\partial f^n}{\partial\lambda_i}(\lambda)\cdot h^n(\lambda) + \frac{\partial g}{\partial\lambda_i}(\lambda)\cdot\frac{h^n(\lambda)}{g(\lambda)}, \qquad \text{for all } \lambda\in[0,1]^N\backslash\{e_\emptyset,e_N\}. \tag{138}$$

Moreover, we obtain the following approximations for all $\lambda\in G_\varepsilon\cap D'(n)$:

$$\frac{\partial f^n}{\partial\lambda_i}(\lambda) = -c(\bar\lambda)\cdot n\cdot\left[\sum_{k\neq i}2\left(\lambda_k-\bar\lambda\right)\cdot -\frac{1}{|N|}+2(\lambda_i-\bar\lambda)\cdot\left(1-\frac{1}{|N|}\right)\right] + \frac{1-2\bar\lambda}{2|N|(\bar\lambda(1-\bar\lambda))^2}\cdot n\cdot\|\lambda-\bar\lambda\cdot e_N\|^2$$

$$
\begin{aligned}
&= -c(\bar{\lambda}) \cdot n \cdot 2\left(\lambda_i - \bar{\lambda}\right) + \frac{1-2\bar{\lambda}}{2|N|(\bar{\lambda}(1-\bar{\lambda}))^2} \cdot n \cdot \|\lambda - \bar{\lambda}\cdot e_N\|^2 \\
&= \mathcal{O}^{\varepsilon}(n^{\frac{1}{2}+\frac{1}{8|N|}}) + \mathcal{O}^{\varepsilon}(n^{\frac{1}{4|N|}}) \qquad (139)\\
&= \mathcal{O}^{\varepsilon}(n^{\frac{1}{2}+\frac{1}{8|N|}}),
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial g}{\partial \lambda_i}(\lambda) &= \mathcal{O}^{\varepsilon}(1), \\
(g(\lambda))^{-1} &= \mathcal{O}(1), \\
h^n(\lambda) &= \mathcal{O}^{\varepsilon}(n^{\frac{1}{2}(1-|N|)}),
\end{aligned}
$$

where (139) follows from $|\lambda_i - \bar{\lambda}| \leq \|\lambda - \bar{\lambda}\cdot e_N\| \leq d'_n = \mathcal{O}(n^{-\frac{1}{2}+\frac{1}{8|N|}})$. Then, the result follows from substituting these equations in (138). $\square$

Lemma D.2 *Let $R \in \mathcal{R}$. Then, we have for all $m \in \{1,\ldots,p\}$ that*

$$
\sum_{\lambda \in G^n_{\varepsilon}\cap D(n)\cap K_m} h^n(\lambda) = n^{|N|} \cdot \sum_{\lambda \in G^n_{\varepsilon}\cap D(n)\cap K_m} \int_{C^n(\lambda)} h^n(\lambda^*)d\lambda^* + \mathcal{O}^{\varepsilon}(n^{\frac{5}{8}}),
$$

where $C^n(\lambda)$ is defined in (93)

Proof Let $\varepsilon > 0$. It is sufficient to show this result for all $n \in \mathbb{N}$ such that $n > \frac{2}{\varepsilon}$. Let $\lambda \in G^n_{\varepsilon} \cap D(n)$. From (109) and (123) follows that

$$
C^n(\lambda) \subset G_{\frac{1}{2}\varepsilon} \cap D'(n). \tag{140}
$$

We get from (140) and Lemma D.1 that h^n is differentiable in λ^* for all $\lambda^* \in C^n(\lambda)$. Applying Taylor's theorem yields that

$$
h^n(\lambda) - h^n(\lambda^*) = \sum_{i\in N} \frac{\partial h}{\partial \lambda_i}(\chi)\cdot(\lambda_i - \lambda^*_i), \text{ for all } \lambda^* \in C^n(\lambda), \text{ where } \chi \in \operatorname{conv}\{\lambda, \lambda^*\}. \tag{141}
$$

Here, as $\chi \in C^n(\lambda)$, we get from Lemma D.1 that

$$
\frac{\partial h}{\partial \lambda_i}(\chi) = \mathcal{O}^{\varepsilon}(n^{-\frac{1}{2}|N|+1+\frac{1}{8|N|}}), \qquad \text{for all } i \in N. \tag{142}
$$

So, as $|\lambda_i - \lambda^*_i| \leq n^{-1}$ for all $\lambda^* \in C^n(\lambda)$ and $i \in N$, we get from (141) and (142) that

$$
\begin{aligned}
h^n(\lambda) - h^n(\lambda^*) &= |N| \cdot \mathcal{O}^{\varepsilon}(n^{-\frac{1}{2}|N|+1+\frac{1}{8|N|}}) \cdot n^{-1} \\
&= \mathcal{O}^{\varepsilon}(n^{-\frac{1}{2}|N|+\frac{1}{8|N|}}),
\end{aligned}
$$

for all $\lambda^* \in C^n(\lambda)$. From this, we directly get

$$
h^n(\lambda) - n^{|N|} \cdot \int_{C^n(\lambda)} h^n(\lambda^*)d\lambda^* = \mathcal{O}^{\varepsilon}(n^{-\frac{1}{2}|N|+\frac{1}{8|N|}}), \qquad \text{for all } \lambda \in G^n_{\varepsilon} \cap D(n). \tag{143}
$$

Moreover, from (112) we get

$$
\begin{aligned}
\sharp(G^n_{\varepsilon} \cap D(n) \cap K_m) &\leq \sharp(G^n_{\varepsilon} \cap D(n)) \\
&= \mathcal{O}(n^{\frac{1}{2}|N|+\frac{5}{8}-\frac{1}{8|N|}}). \qquad (144)
\end{aligned}
$$

Hence, from (143) and (144) follows that

$$
\left| \sum_{\lambda \in G^n_{\varepsilon}\cap D(n)\cap K_m} \left(h^n(\lambda) - n^{|N|} \cdot \int_{C^n(\lambda)} h^n(\lambda^*)d\lambda^* \right) \right| \leq \sharp(G^n_{\varepsilon} \cap D(n) \cap K_m) \cdot \mathcal{O}^{\varepsilon}(n^{-\frac{1}{2}|N|+\frac{1}{8|N|}})
$$

$$= \mathcal{O}^\varepsilon(n^{\frac{5}{8}}).$$

This concludes the result. □

Lemma D.3 *Let $R \in \mathcal{R}$. Then, we have for all $m \in \{1, \ldots, p\}$ that*

$$\sum_{\lambda^* \in G_\varepsilon^n \cap D(n) \cap K_m} \int_{C^n(\lambda^*)} h^n(\lambda) d\lambda = \int_{G_\varepsilon \cap D(n) \cap K_m} h^n(\lambda) d\lambda + \mathcal{O}^\varepsilon(n^{-|N|+\frac{5}{8}}).$$

Proof Let $\varepsilon > 0$ and define $D''(n) = D^{d''_n}$, where $d''_n = d_n - (\sqrt{|N|}/n)$. It is sufficient to show this result for all $n \in \mathbb{N}$ such that $n > \frac{2}{\varepsilon}$. Define $A = \bigcup_{\lambda^* \in G_\varepsilon^n \cap D(n) \cap K_m} C^n(\lambda^*)$ and $B = G_\varepsilon \cap D(n) \cap K_m$. Moreover, define

$$\begin{aligned}
E_1^n &= B(n/(\sqrt{|N|}+1)) \cap G_{\frac{1}{2}\varepsilon} \\
E_2^n &= [G_{\varepsilon-(1/n)} \cap D'(n)] \backslash D''(n) \\
E_3^n &= [D'(n) \cap G_{\varepsilon-(1/n)}] \backslash G_{\varepsilon+(1/n)},
\end{aligned}$$

where the set $B(n)$ is defined in (97). We first show

$$(A \backslash B) \cup (B \backslash A) \subset E_1^n \cup E_2^n \cup E_3^n. \tag{145}$$

Let $y_1 \in A \backslash B$, so we have $y_1 \in C^n(\lambda)$ for some $\lambda \in G_\varepsilon^n \cap D(n) \cap K_m$. If $y_1 \notin K_m$, there is a $\lambda' \in [0,1]^N \backslash L(R)$ such that $\lambda' \in \text{conv}\{\lambda, y_1\}$ and, so, $y_1 \in E_1^n$. If $y_1 \notin D(n)$, we have according to (108) that $\|y_1 - \bar{y}_1 \cdot e_N\| < (\sqrt{|N|}/n) + d_n = d'_n$ and, so, $y_1 \in E_2^n$. If $y_1 \notin G_\varepsilon^n$, then $\bar{y}_1 < \varepsilon$ or $\bar{y}_1 > 1 - \varepsilon$ and hence we have according to (123) that $\varepsilon - (1/n) \leq \bar{y}_1 \leq 1 - (\varepsilon - (1/n))$ and, so, $y_1 \in E_3^n$. Now, let $y_2 \in B \backslash A$, so we have $y_2 \in G_\varepsilon \cap D(n) \cap K_m$ and there does not exist a $\lambda \in G_\varepsilon^n \cap D(n) \cap K_m$ such that $y_2 \in C^n(\lambda)$. Let λ such that $y_2 \in C^n(\lambda)$. If $\lambda \notin K_m$, there exists an $\lambda' \in [0,1]^N \backslash L(R)$ such that $\lambda' \in \text{conv}\{\lambda, y_2\}$ and, so, $y_2 \in E_1^n$. If $\lambda \notin D(n)$, we get from the triangle inequality that $\|y_2 - \bar{y}_2 \cdot e_N\| \geq \|\lambda - \bar{\lambda} \cdot e_N\| - \|y_2 - \lambda\| \geq d_n - (\sqrt{|N|}/n) = d''_n$ and, so, $y_2 \notin D''(n)$. So, $y_2 \in E_2^n$. If $\lambda \notin G_\varepsilon^n$, then $\bar{\lambda} < \varepsilon$ or $\bar{\lambda} > 1 - \varepsilon$ and hence $\bar{y}_2 = \bar{\lambda} + (\bar{y}_2 - \bar{\lambda}) < \varepsilon + (1/n)$ or $\bar{y}_2 < 1 - (\varepsilon + (1/n))$ and so, $y_2 \notin G_{\varepsilon+(1/n)}$. So, $y_2 \in E_3^n$. Hence, we have shown (145). Then, we get

$$\begin{aligned}
\left| \int_A h^n(\lambda) d\lambda - \int_B h^n(\lambda) d\lambda \right| &\leq \int_{A \backslash B} h^n(\lambda) d\lambda + \int_{B \backslash A} h^n(\lambda) d\lambda & (146) \\
&\leq \int_{E_1^n \cup E_2^n \cup E_3^n} h^n(\lambda) d\lambda & (147) \\
&\leq \sum_{k=1}^3 \int_{E_k^n} h^n(\lambda) d\lambda & (148) \\
&\leq \sum_{k=1}^3 \nu(E_k^n) \cdot \mathcal{O}^\varepsilon(n^{\frac{1}{2}(1-|N|)}) & (149) \\
&= \mathcal{O}^\varepsilon(n^{-|N|+\frac{5}{8}}). & (150)
\end{aligned}$$

Here, (146) follows from $\int_A h^n(\lambda) d\lambda - \int_B h^n(\lambda) d\lambda = \int_{A \backslash B} h^n(\lambda) d\lambda - \int_{B \backslash A} h^n(\lambda) d\lambda$, (147) follows from (145), (148) is a standard rule of integration, (149) follows from $h^n(\lambda) = \mathcal{O}^\varepsilon(n^{\frac{1}{2}(1-|N|)})$ for all $\lambda \in G_{\frac{1}{2}\varepsilon}$ and (150) follows from $\nu(E_1^n) = \mathcal{O}^\varepsilon(n^{-\frac{1}{2}|N|+\frac{1}{8}})$ (see (122)) and we get in a similar fashion as for (122) via a Gram-Schmidt process that

$$\begin{aligned}
\nu(E_2^n) &= \mathcal{O}\left(\left(n^{-\frac{1}{2}+\frac{1}{8|N|}} + (\sqrt{|N|}/n) \right)^{|N|-1} - \left(n^{-\frac{1}{2}+\frac{1}{8|N|}} - (\sqrt{|N|}/n) \right)^{|N|-1} \right) \\
&= \mathcal{O}^\varepsilon(n^{-\frac{1}{2}|N|+\frac{1}{8}}), \\
\nu(E_3^n) &= \mathcal{O}((n^{-\frac{1}{2}+\frac{1}{8|N|}} + (\sqrt{|N|}/n))^{|N|-1} \cdot n^{-1}) \\
&= \mathcal{O}^\varepsilon(n^{-\frac{1}{2}|N|-\frac{3}{8}}).
\end{aligned}$$

This concludes the proof. □

Lemma D.4 *For all $t \in (0,1)$ it holds that*

$$\int_0^{n^{\frac{1}{4|N|}}/2t(1-t)} e^{-s} \cdot s^{\frac{1}{2}(|N|-3)} ds = \Gamma\left(\frac{1}{2}|N| - \frac{1}{2}\right) + \mathcal{O}(n^{-\infty}),$$

where Γ is the Gamma function:

$$\Gamma(\kappa) = \int_0^\infty e^{-t} \cdot t^{\kappa-1} dt, \qquad \textit{for all } \kappa > 0. \tag{151}$$

Proof We get

$$\begin{aligned}
\Gamma\left(\frac{1}{2}|N| - \frac{1}{2}\right) - \int_0^{n^{\frac{1}{4|N|}}/2t(1-t)} e^{-s} \cdot s^{\frac{1}{2}(|N|-3)} ds &= \int_{n^{\frac{1}{4|N|}}/2t(1-t)}^\infty e^{-s} \cdot s^{\frac{1}{2}(|N|-3)} ds && (152)\\
&\le K \cdot \int_{n^{\frac{1}{4|N|}}/2t(1-t)}^\infty e^{-\frac{1}{2}s} ds && (153)\\
&= K \cdot 2e^{-n^{\frac{1}{4|N|}}/4t(1-t)} && (154)\\
&\le K \cdot 2e^{-n^{\frac{1}{4|N|}}} && (155)\\
&= \mathcal{O}(n^{-\infty}), && (156)
\end{aligned}$$

where $K > 0$. Here, (152) is a standard integration rule, (153) follows from that there exists a constant $K > 0$ such that $e^{-s} \cdot s^{\frac{1}{2}(|N|-3)} < K \cdot e^{-\frac{1}{2}s}$ for all $s > 1$, (154) follows from $\int_a^b e^{-\frac{1}{2}s} ds = -2(e^{-\frac{1}{2}b} - e^{-\frac{1}{2}a})$ for all $a \le b$, (155) follows from $4t(1-t) \le 1$ for all $t \in (0,1)$ and (156) follows from the fact that $(e^{-1})^{n^{\frac{1}{4|N|}}} = \mathcal{O}(n^{-\infty})$. This concludes the proof. □

Proof of Proposition 6.6: We get

$$\begin{aligned}
K_i^n(R) &= \sum_{m=1}^{p^*} E_{\mathbb{Q}_m}[X_i] \cdot \frac{1}{n} \cdot \sum_{\lambda \in G_\varepsilon^n \cap D(n) \cap K_m} h^n(\lambda) + \mathcal{O}(\varepsilon) + \mathcal{O}^\varepsilon(n^{-\frac{1}{4}}) && (157)\\
&= \sum_{m=1}^{p^*} E_{\mathbb{Q}_m}[X_i] \cdot n^{|N|-1} \cdot \sum_{\lambda \in G_\varepsilon^n \cap D(n) \cap K_m} \int_{C^n(\lambda)} h^n(\lambda^*) d\lambda^* + \mathcal{O}(\varepsilon) + \mathcal{O}^\varepsilon(n^{-\frac{1}{4}}) && (158)\\
&= \sum_{m=1}^{p^*} E_{\mathbb{Q}_m}[X_i] \cdot n^{|N|-1} \cdot \int_{G_\varepsilon \cap D(n) \cap K_m} h^n(\lambda) d\lambda + \mathcal{O}(\varepsilon) + \mathcal{O}^\varepsilon(n^{-\frac{1}{4}}) && (159)\\
&= \sum_{m=1}^{p^*} E_{\mathbb{Q}_m}[X_i] \cdot n^{\frac{1}{2}(|N|-1)} \cdot (2\pi)^{\frac{1}{2}(1-|N|)} \cdot |N|^{-\frac{1}{2}} && (160)\\
&\quad \cdot \int_{G_\varepsilon \cap D(n) \cap K_m} \left(e^{-\frac{1}{2\bar\lambda(1-\bar\lambda)} \cdot n \cdot \|\lambda - \bar\lambda \cdot e_N\|^2}\right) \cdot (\bar\lambda(1-\bar\lambda))^{\frac{1}{2}(1-|N|)} d\lambda + \mathcal{O}(\varepsilon) + \mathcal{O}^\varepsilon(n^{-\frac{1}{4}})\\
&= \sum_{m=1}^{p^*} E_{\mathbb{Q}_m}[X_i] \cdot n^{\frac{1}{2}(|N|-1)} \cdot (2\pi)^{\frac{1}{2}(1-|N|)} && (161)\\
&\quad \cdot \int_\varepsilon^{1-\varepsilon} \int_0^{d_n} \int_{S_m} e^{-\frac{1}{2t(1-t)} r^2 n} \cdot (t(1-t))^{\frac{1}{2}(1-|N|)} \cdot r^{|N|-2} d\omega dr dt + \mathcal{O}(\varepsilon) + \mathcal{O}^\varepsilon(n^{-\frac{1}{4}})\\
&= \sum_{m=1}^{p^*} E_{\mathbb{Q}_m}[X_i] \cdot n^{\frac{1}{2}(|N|-1)} \cdot (2\pi)^{\frac{1}{2}(1-|N|)} \cdot \mu(S_m) && (162)
\end{aligned}$$

$$
\begin{aligned}
&\cdot \int_{\varepsilon}^{1-\varepsilon} \int_{0}^{d_n} e^{-\frac{1}{2t(1-t)} r^2 n} \cdot (t(1-t))^{\frac{1}{2}(1-|N|)} \cdot r^{|N|-2} dr dt + \mathcal{O}(\varepsilon) + \mathcal{O}^{\varepsilon}(n^{-\frac{1}{4}}) \\
&= \sum_{m=1}^{p^*} E_{\mathbb{Q}_m}[X_i] \cdot n^{\frac{1}{2}(|N|-1)} \cdot (2\pi)^{\frac{1}{2}(1-|N|)} \cdot \phi_m \cdot 2 \frac{\pi^{-\frac{1}{2}(1-|N|)}}{\Gamma\left(\frac{1}{2}|N| - \frac{1}{2}\right)} && (163) \\
&\cdot \int_{\varepsilon}^{1-\varepsilon} \int_{0}^{d_n} e^{-\frac{1}{2t(1-t)} r^2 n} \cdot (t(1-t))^{\frac{1}{2}(1-|N|)} \cdot r^{|N|-2} dr dt + \mathcal{O}(\varepsilon) + \mathcal{O}^{\varepsilon}(n^{-\frac{1}{4}}) \\
&= \sum_{m=1}^{p^*} E_{\mathbb{Q}_m}[X_i] \cdot \phi_m \cdot n^{\frac{1}{2}(|N|-1)} \cdot 2^{1\frac{1}{2}-\frac{1}{2}|N|} \cdot \frac{1}{\Gamma\left(\frac{1}{2}|N| - \frac{1}{2}\right)} \cdot \left(\frac{2}{n}\right)^{\frac{1}{2}(|N|-2)} && (164) \\
&\cdot \int_{\varepsilon}^{1-\varepsilon} \int_{0}^{n^{\frac{1}{4|N|}}/2t(1-t)} e^{-s} \cdot s^{\frac{1}{2}(|N|-2)} \cdot (t(1-t))^{-\frac{1}{2}} \cdot \sqrt{\frac{t(1-t)}{2ns}} ds dt + \mathcal{O}(\varepsilon) + \mathcal{O}^{\varepsilon}(n^{-\frac{1}{4}}) \\
&= \sum_{m=1}^{p^*} E_{\mathbb{Q}_m}[X_i] \cdot \phi_m \cdot \cdot \frac{1}{\Gamma\left(\frac{1}{2}|N| - \frac{1}{2}\right)} && (165) \\
&\cdot \int_{\varepsilon}^{1-\varepsilon} \int_{0}^{n^{\frac{1}{4|N|}}/2t(1-t)} e^{-s} \cdot s^{\frac{1}{2}(|N|-3)} ds dt + \mathcal{O}(\varepsilon) + \mathcal{O}^{\varepsilon}(n^{-\frac{1}{4}}) \\
&= \sum_{m=1}^{p^*} E_{\mathbb{Q}_m}[X_i] \cdot \phi_m \cdot \frac{1}{\Gamma\left(\frac{1}{2}|N| - \frac{1}{2}\right)} \cdot \int_{\varepsilon}^{1-\varepsilon} \left(\Gamma\left(\frac{1}{2}|N| - \frac{1}{2}\right) + \mathcal{O}(n^{-\infty})\right) dt + \mathcal{O}(\varepsilon) + \mathcal{O}^{\varepsilon}(n^{-\frac{1}{4}}) && (166) \\
&= \sum_{m=1}^{p^*} \phi_m \cdot E_{\mathbb{Q}_m}[X_i] + \mathcal{O}(\varepsilon) + \mathcal{O}^{\varepsilon}(n^{-\frac{1}{4}}) && (167)
\end{aligned}
$$

Here, (157) follows from Proposition 6.5, (158) follows from Lemma D.2 and (159) follows from Lemma D.3, (160) follows from Definition 6.4, (161) follows from the polar coordinate transformation $\lambda = t \cdot e_N + r \cdot \omega$ and $d\lambda = r^{|N|-2} \cdot |N|^{\frac{1}{2}} d(t, r, \omega)$, (162) follows from the fact that $\int_{S_m} d\omega = \mu(S_m)$, (163) follows from the well-known result

$$
\mu(S_m) = \phi_m \cdot \mu(S) = \phi_m \cdot 2 \frac{\pi^{-\frac{1}{2}(1-|N|)}}{\Gamma\left(\frac{1}{2}|N| - \frac{1}{2}\right)},
$$

where Γ is defined in (151), (164) follows from the transformation $s = \frac{r^2 n}{2t(1-t)}$ and $dr = \sqrt{\frac{t(1-t)}{2ns}} ds$, (165) follows from canceling of some terms and (166) follows from Lemma D.4. This concludes the proof.